

GenCore version 5.1.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.  
 OM protein - protein search, using sw model  
 Run on: September 21, 2004, 13:35:22 ; Search time 88.4672 Seconds  
 (without alignments)  
 1194.261 Million cell updates/sec

Title: US-10-037-860-4  
 Perfect score: 1729  
 Sequence: 1 MAMTLLDWCRCMDVNSQRT.....LTGAGEGPGPKPLSVAGADP 329

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
 8: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
 12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
 19: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1729	100.0	329	13	US-10-037-860-4
2	1666	96.4	318	12	US-09-804-014A-40
3	1602	92.7	353	9	US-09-965-529-7
4	1602	92.7	353	10	US-09-969-680A-7
5	887.5	51.3	321	12	US-09-804-014A-39
6	887.5	51.3	351	9	US-09-965-529-1
7	887.5	51.3	351	10	US-09-969-680A-1
8	887.5	51.3	351	12	US-09-804-014A-16
9	887.5	51.3	351	15	US-10-341-434-10
10	874.5	50.6	312	12	US-09-804-014A-73
11	874.5	50.6	312	12	US-09-804-014A-73
12	766.5	44.3	463	13	US-10-037-860-13
13	744	43.0	452	16	US-10-408-765A-2385
14	620	35.9	299	15	US-10-094-749-1978
15	564	32.6	283	13	US-10-037-860-11

Sequence 7, Appli  
 Sequence 41, Appl  
 Sequence 42, Appl  
 Sequence 38, Appl  
 Sequence 1208, Ap  
 Sequence 33747, A  
 Sequence 9, Appli  
 Sequence 34645, A  
 Sequence 2992, Ap  
 Sequence 355, App  
 Sequence 355, App  
 Sequence 45, Appl  
 Sequence 49641, A  
 Sequence 50770, A  
 Sequence 65151, A  
 Sequence 66072, A  
 Sequence 17979, A  
 Sequence 47973, A  
 Sequence 5485, Ap  
 Sequence 58461, A  
 Sequence 25, Appl  
 Sequence 114, App  
 Sequence 144559, A  
 Sequence 50904, A  
 Sequence 1013, Ap  
 Sequence 11, Appl  
 Sequence 62912, A  
 Sequence 2, Appli  
 Sequence 10509, A  
 Sequence 51122, A

16 441 25.5 195 13 US-10-037-860-7  
 17 353.5 20.4 120 12 US-09-804-014A-41  
 18 338.5 19.6 120 12 US-09-804-014A-42  
 19 335.5 19.4 403 15 US-10-094-466-38  
 20 326 18.9 337 12 US-10-296-115-1208  
 21 304 17.6 204 14 US-10-029-386-33747  
 22 256.5 14.8 149 13 US-10-037-860-9  
 23 246.5 14.3 116 9 US-09-864-761-34645  
 24 192 11.1 538 16 US-10-408-765A-2992  
 25 133 7.7 584 12 US-10-221-278-355  
 26 133 7.7 584 15 US-10-291-172-355  
 27 120.5 7.0 5245 14 US-10-329-079-45  
 28 115.5 6.7 555 12 US-10-282-122A-49641  
 29 114.5 6.6 555 12 US-10-282-122A-50770  
 30 114 6.6 558 12 US-10-282-122A-65151  
 31 111 6.4 558 12 US-10-282-122A-66072  
 32 109 6.3 531 15 US-10-369-493-17979  
 33 107.5 6.2 526 12 US-10-282-122A-47973  
 34 106.5 6.2 503 9 US-09-738-626-5485  
 35 104.5 6.0 556 12 US-10-282-122A-58461  
 36 103 6.0 935 14 US-10-080-608A-25  
 37 103 6.0 935 15 US-10-370-685-114  
 38 103 6.0 1855 16 US-10-437-963-144559  
 39 101.5 5.9 495 12 US-10-425-114-50904  
 40 99.5 5.8 1997 16 US-10-408-765A-1013  
 41 99.5 5.8 5245 14 US-10-329-079-11  
 42 98 5.7 553 12 US-10-282-122A-62912  
 43 98 5.7 1009 16 US-10-607-631-2  
 44 98 5.7 1276 14 US-10-156-761-10509  
 45 97 5.6 538 12 US-10-282-122A-51122

## ALIGNMENTS

RESULT 1  
 US-10-037-860-4  
 ; Sequence 4, Application US/10037860  
 ; Publication No. US20020123114A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jerome B. Posner  
 ; APPLICANT: Joseph O. Dalmau  
 ; APPLICANT: Myrna R. Rosenfeld  
 ; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA  
 ; TITLE OF INVENTION: ANTIBODIES  
 ; FILE REFERENCE: 2581.1004-004  
 ; CURRENT APPLICATION NUMBER: US/10/037,860  
 ; CURRENT FILING DATE: 2001-01-04  
 ; PRIOR APPLICATION NUMBER: 09/189,527  
 ; PRIOR FILING DATE: 1998-11-10  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 329  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 ; US-10-037-860-4

Query Match 100.0%; Score 1729; DB 13; Length 329;  
 Best Local Similarity 100.0%; Pred. No. 5e-166;  
 Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGPVNCDAEIEETLQAAMPQVSYRMLGRMFWRN 60  
 1 MAMTLLDWCRCMDVNSQRTLLVWGPVNCDAEIEETLQAAMPQVSYRMLGRMFWRN 60

QY 61 AKAAELLELTGAVDYAAIIPREMFGKGVKWLFPKPTSDAEFLERLHLFLAREGWTVDYA 120  
 61 AKAAELLELTGAVDYAAIIPREMFGKGVKWLFPKPTSDAEFLERLHLFLAREGWTVDYA 120

QY 121 RVLFQNPPTTGPMPAPMLNYILDNIQPLVSIWYKRLTLFSGKGHPRAWRFNDFDW 180  
 121 RVLFQNPPTTGPMPAPMLNYILDNIQPLVSIWYKRLTLFSGKGHPRAWRFNDFDW 180

Db 181 LEHTNEVLEWQSDVDEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240  
Qy 241 ESSRDAQIKFLNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
Db 241 ESSRDAQIKFLNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
Qy 301 SGAIRROLWLTGAGEGPG 318  
Db 301 SGAIRROLWLTGAGEGPG 318  
RESULT 3  
US-09-965-529-7  
; Sequence 7, Application US/09965529  
; Publication No. US20020182671A1  
; GENERAL INFORMATION:  
; APPLICANT: LAL, Preeti  
; APPLICANT: YUE, Henry  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: BURFORD, Neil  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: BAUGHN, Maria R.  
; APPLICANT: LU, Dzung Aina M.  
; APPLICANT: PATTERSON, Chandra  
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0731 USA  
; CURRENT APPLICATION NUMBER: US/09/965,529  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315  
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: PERL Program  
; SEQ ID NO 7  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020182671A1 2483172CD1  
US-09-965-529-7  
Query Match 92.7%; Score 1602; DB 9; Length 353;  
Best Local Similarity 96.6%; Pred. No. 3.9e-153;  
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
Qy 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAAAMPQVSYRMLGRMFREEN 60  
Db 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAAAMPQVSYRMLGRMFREEN 60  
Qy 61 AKAAELLELTGAVDYAAIPREMPKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
Db 61 AKAAELLELTGAVDYAAIPREMPKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
Qy 121 RVLGFGQNPPTPGPEMPAEMNLNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPW 180  
Db 121 RVLGFGQNPPTPGPEMPAEMNLNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPW 180  
Qy 181 LEHTNEVLEWQSDVDEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240  
Db 181 LEHTNEVLEWQSDVDEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240  
Qy 241 ESSRDAQIKFLNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
Db 241 ESSRDAQIKFLNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
Qy 301 SGAIRROLWLTGAGEGPG 319  
Db 301 SGAIRROLWLTGAGEGPG 319  
RESULT 4  
US-09-969-680A-7

Qy 181 LEHTNEVLEWQSDVDEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240  
Db 181 LEHTNEVLEWQSDVDEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240  
Qy 241 ESSRDAQIKFLNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
Db 241 ESSRDAQIKFLNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
Qy 301 SGAIRROLWLTGAGEGPGPKPLSVAGADP 329  
Db 301 SGAIRROLWLTGAGEGPGPKPLSVAGADP 329  
RESULT 2  
US-09-804-014A-40  
; Sequence 40, Application US/09804014A  
; Publication No. US20030064489A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Li  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Vernet, Corine  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Majumder, Kumud  
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-721 US  
; CURRENT APPLICATION NUMBER: US/09/804,014A  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: 60/188,316  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 60/188,277  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 60/189,139  
; PRIOR FILING DATE: 2000-03-14  
; PRIOR APPLICATION NUMBER: 60/189,140  
; PRIOR FILING DATE: 2000-03-14  
; PRIOR APPLICATION NUMBER: 60/190,401  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/190,231  
; PRIOR FILING DATE: 2000-03-17  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 40  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: VARIANT  
; LOCATION: (20)  
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the  
; OTHER INFORMATION: specification  
US-09-804-014A-40  
Query Match 96.4%; Score 1666; DB 12; Length 318;  
Best Local Similarity 99.7%; Pred. No. 1.1e-159;  
Matches 317; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAAAMPQVSYRMLGRMFREEN 60  
Db 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAAAMPQVSYRMLGRMFREEN 60  
Qy 61 AKAAELLELTGAVDYAAIPREMPKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
Db 61 AKAAELLELTGAVDYAAIPREMPKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
Qy 121 RVLGFGQNPPTPGPEMPAEMNLNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPW 180  
Db 121 RVLGFGQNPPTPGPEMPAEMNLNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPW 180  
Qy 181 LEHTNEVLEWQSDVDEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240

```
; Sequence 7, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/149,641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164,203
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030124649A1 2483172CD1
US-09-969-680A-7

Query Match          92.7%; Score 1602; DB 10; Length 353;
Best Local Similarity 96.6%; Pred. No. 3.9e-153;
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAMQVSVYMLGRMFWEEN 60
Db 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAMQVSVYMLGRMFWEEN 60
QY 61 AKAALELTGAVDYAAIPREMPGKGGVWVLFKPPTSDAEFLERLHLFLAREGWTVDVA 120
Db 61 AKAALELTGAVDYAAIPREMPGKGGVWVLFKPPTSDAEFLERLHLFLAREGWTVDVA 120
QY 121 RVLGQNPETTPGPEPAEMNLVDNIQPLVESIWKYKLTLPFGKHGHPRAWGNFDPW 180
Db 121 RVLGQNPETTPGPEPAEMNLVDNIQPLVESIWKYKLTLPFGKHGHPRAWGNFDPW 180
QY 181 LEHNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQVFGSV 240
Db 181 LEHNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQVFGSV 240
QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAIKDNVNNQARLEQVIAGANH 300
Db 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAIKDNVNNQARLEQVIAGANH 300
QY 301 SGAIRQLMLTGAGEGPG 319
Db 301 SGAIRQLMLTGAGEGPG 319

RESULT 5
US-09-804-014A-39
; Sequence 39, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
```

```
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 39
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-39
```

```
Query Match          51.3%; Score 887.5; DB 12; Length 321;
Best Local Similarity 55.4%; Pred. No. 7.1e-81;
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAMQVSVYMLGRMFWEEN 59
Db 1 MTELLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAMQVSVYMLGRMFWEEN 60
QY 60 NAKAALELTGAVDYAAIPREMPGKGGVWVLFKPPTSDAEFLERLHLFLAREGWTVDV 119
Db 61 NRKVALVGLTAETSHALVPEIKPGGIGVIRVIFKPPDPDNTFJSLRNEFLAGEGMYGEL 120
QY 120 ARVLGFQNPETTPG--PEMPAEMNLVDNIQPLVESIWKYKLTLPFGKHGHPRAWRG 175
Db 121 SRALGHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLVFSGRESPEFGE 179
QY 176 NFDPWLEHTEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 235
Db 180 EFGRWPFHTTQMIRKAWQVDPVEKRRRLMESLRGPAADVIRILKSNPAITTVDECLALEE 239
QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAIKDNVNNQARLEQVI 295
Db 240 VFGTUNPRLQVLYLTYYQDEEKLSAYVIRLEPLLOKVEKGAIKDNVNNQARLEQVI 299
QY 296 AGAHSGAIRQLMLTGAGEGPG 318
Db 300 AGAVHK-TIRRELNLDPDGPAG 321
```

```
RESULT 6
US-09-965-529-1
; Sequence 1, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 1
```



```

; 121 SRALCHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLVFSGRESPEEGEE 179
; 176 NFDPLWHTNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPNPAITTAECCLKALEQ 235
; 180 EFGRMWFHTTQMIKAWQVDPVEKRRRLLESURGPALDVIRVLKINNPLIITVDECLQALEE 239
; 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKQVVEKGAKDKNNVQARLEQVI 295
; 240 VFGVTDPNPRELQVKYLTITTYQKDEEKL SAYVLRLEPLLOKLVQVGAERDAVNAQRLDQVI 299
; 296 AGANHSAGAIRRQLWLTGAGEGPG 318
; 300 AGAVHK-TIRRELNPEDGPAPG 321

```

## RESULT 9

```

US-10-341-434-10
; Sequence 10, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-10

```

```

Query Match 51.3%; Score 887.5; DB 15; Length 351;
Best Local Similarity 55.4%; Pred. No. 8.1e-81;
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;

```

```

QY 1 MAMTLLEDWCRCMDVNSQRTLLVWGIPVNCDEABIEETLQAA-POVSYRMLGRMFWRREE 59
Db 1 MTLRLLEDWCRCMDMNPCKALLIAGISQSCSVABIEEALQAGLAPLGEYRLLGEMFRDE 60
QY 60 NAKAALLELTGADVAAIPREMPCKGGVWKVLFKPTSDAEFLERHLFLAREGWTVDV 119
Db 61 NRKVALVGLTAETSHALVPKEIPGKGIMRWIFKPPDPDNTFLSRLNEFLAGEGMTVGEL 120
QY 120 ARVLGFQNP-PTPG--PMPAEMLNVIIDNVIOPLVESIWKRLTLFSGKGHPRAWRG 175
Db 121 SRALCHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLVFSGRESPEEGEE 179
QY 176 NFDPLWHTNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPNPAITTAECCLKALEQ 235
Db 180 EFGRMWFHTTQMIKAWQVDPVEKRRRLLESURGPALDVIRVLKINNPLIITVDECLQALEE 239
QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKQVVEKGAKDKNNVQARLEQVI 295
Db 240 VFGVTDPNPRELQVKYLTITTYQKDEEKL SAYVLRLEPLLOKLVQVGAERDAVNAQRLDQVI 299
QY 296 AGANHSAGAIRRQLWLTGAGEGPG 318
Db 300 AGAVHK-TIRRELNPEDGPAPG 321

```

## RESULT 10

```

US-09-804-014A-73
; Sequence 73, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li

```

```

; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-73

```

```

Query Match 50.6%; Score 874.5; DB 12; Length 312;
Best Local Similarity 55.9%; Pred. No. 1.4e-79;
Matches 175; Conservative 53; Mismatches 78; Indels 7; Gaps 5;

```

```

QY 1 MAMTLLEDWCRCMDVNSQRTLLVWGIPVNCDEABIEETLQAA-POVSYRMLGRMFWRREE 59
Db 1 MTLRLLEDWCRCMDMNPCKALLIAGISQSCSVABIEEALQAGLAPLGEYRLLGEMFRDE 60
QY 60 NAKAALLELTGADVAAIPREMPCKGGVWKVLFKPTSDAEFLERHLFLAREGWTVDV 119
Db 61 NRKVALVGLTAETSHALVPKEIPGKGIMRWIFKPPDPDNTFLSRLNEFLAGEGMTVGEL 120
QY 120 ARVLGFQNP-PTPG--PMPAEMLNVIIDNVIOPLVESIWKRLTLFSGKGHPRAWRG 175
Db 121 SRALCHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLVFSGRESPEEGEE 179
QY 176 NFDPLWHTNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPNPAITTAECCLKALEQ 235
Db 180 EFGRMWFHTTQMIKAWQVDPVEKRRRLLESURGPALDVIRVLKINNPLIITVDECLQALEE 239
QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKQVVEKGAKDKNNVQARLEQVI 295
Db 240 VFGVTDPNPRELQVKYLTITTYQKDEEKL SAYVLRLEPLLOKLVQVGAERDAVNAQRLDQVI 299
QY 296 AGANHSAGAIRRQL 308
Db 300 AGAVHK-TIRREL 311

```

## RESULT 11

```

US-09-804-014A-74
; Sequence 74, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US

```

```

; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 74
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-74

Query Match      50.6%; Score 874.5; DB 12; Length 312;
Best Local Similarity 55.9%; Pred. No. 1.4e-79;
Matches 175; Conservative 53; Mismatches 78; Indels 7; Gaps 5;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAAMPQVS-YRMLGRMFWRREE 59
DB 1 MTLRLLEDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAAMPQVS-YRMLGRMFWRREE 60

QY 60 NAKAALLELTGAVDYAAIIPREMPGKGGVWVLFKPPSTDABFLERLHLFLAREGWTVDV 119
DB 61 NKKVALVGLTARTSHALVPKEIPGGGIWRVIFKPPDPDNTFLSLNEFLAGEGTVGEL 120

QY 120 ARVLGFQNPFT--PTFG--PMPAEMLNLYLDNVIQPLVESIWKRLTLPFGKHGPRANRG 175
DB 121 SFALCHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLRVFSGRESPEGEE 179

QY 176 NPDPLWHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQV 235
DB 180 EFRWWEFTTQMIKAWQVDPVDEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQV 239

QY 236 VFGSVSSRDAQIKFLNTYQNGEKL SAYVIRLEPLIQVKEGADIDKDNVQARLEQVI 295
DB 240 VFGVTDNPRELQVLYTTQKDEEKL SAYVIRLEPLIQVKEGADIDKDNVQARLEQVI 299

QY 296 AGANHSGAIRROL 308
DB 300 AGAVHK-TIRREL 311

RESULT 12
US-10-037-860-13
; Sequence 13, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Joseph O. Dalmat
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 463
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-13
```

```

Query Match      44.3%; Score 766.5; DB 13; Length 463;
Best Local Similarity 50.2%; Pred. No. 2.2e-68;
Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAAMPQVS-YRMLGRMFWRREE 59
DB 1 MTLRLLEDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAAMPQVS-YRMLGRMFWRREE 60

QY 60 NAKAALLELTGAVDYAAIIPREMPGKGGVWVLFKPPSTDABFLERLHLFLAREGWTVDV 119
DB 61 NAKAALLELTGAVDYAAIIPREMPGKGGVWVLFKPPSTDABFLERLHLFLAREGWTVDV 120

QY 120 ARVLGFQNPFTPGPEMPAEMLNLY--ILDNVIQPLVESIWKRLTLPFGKHGPRANRG 177
DB 121 NRVLGSDTNCAPRTVISPFEFTWQAQTLGAAVQPLLEQMLYRELRFVSGTISIPGALAF 180

QY 178 DPWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQV 237
DB 181 DAWLEHTTLMQWQVPEGEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQV 240

QY 238 GSVESSRDAQIKFLNTYQNGEKL SAYVIRLEPLIQVKEGADIDKDNVQARLEQVI 297
DB 241 GPVESHKIAQVLCCKAYQEAQEKYSSVIRLEPLIQVKEGADIDKDNVQARLEQVI 300

QY 298 ANHSGAIRROL 310
DB 301 ATLPDKLRDKL 313

RESULT 13
US-10-408-765A-2385
; Sequence 2385, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2385
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2385

Query Match      43.0%; Score 744; DB 16; Length 452;
Best Local Similarity 46.0%; Pred. No. 3.9e-66;
Matches 143; Conservative 64; Mismatches 100; Indels 4; Gaps 2;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAAMPQVS-YRMLGRMFWRREE 59
DB 5 MALTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAAMPQVS-YRMLGRMFWRREE 64

QY 60 NAKAALLELTGAVDYAAIIPREMPGKGGVWVLFKPPSTDABFLERLHLFLAREGWTVDV 119
DB 65 NAKAALLELTGAVDYAAIIPREMPGKGGVWVLFKPPSTDABFLERLHLFLAREGWTVDV 124

QY 120 ARVLGFQNPFTPGPEMPAEMLNLYLDNVIQPLVESIWKRLTLPFGKHGPRANRG 179
DB 125 ARALGC---CSLPAESLDAEVMPQVSPPEPPKSNMYKRLVFSGTASPGSEETFD 181

QY 180 WLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQV 239
```

```
Db 182 WLEQVTEIMPVSEVEKRRRLSLRGPALSIIMRVLQANNDISITVEQCLDALKQIFGD 241
QY 240 VESSRDAQIKFLNTYONPGEKLSAVVIRLEPLLOKVVEKGAIDKDNVNOARLEQVIAGAN 299
Db 242 KEDFRASQRFLOTSPKIGERKSTFLRLLEPLLOKAVHKPLSVNRSTOMIRUKHLLARVA 301
QY 300 HSGAIRRQIWL 310
Db 302 MTPALRGKLEL 312

RESULT 14
US-10-094-749-1978
; Sequence 1978, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1978
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-1978

Query Match 35.9%; Score 620; DB 15; Length 399;
Best Local Similarity 40.9%; Pred. No. 1.2e-53;
Matches 137; Conservative 64; Mismatches 106; Indels 28; Gaps 6;

QY 1 MAMTLEDRCRGMVNSQRTLLVWGPVNCDEAFIEETLQAM-POVSYRMLGRMFWRREE 59
Db 1 MAVTMLQDWCKWGVNARRGLLIGIPEDCDDAEFQESLEAALRPMGHFTVLGKAFREED 60
QY 60 NAKAALLELTGAVDYAAIPREMKGKGVKVLFPKPTSDAEF--LERLHLFLAREGTVQ 117
Db 61 NATAALVELDRVNYALVPREIFETGTGGMNVVFPVRCGSEEFGLGRVFFHFEQEGQWVE 120
QY 118 DVARVLGFQNPFTPTPGMPAEMLNYYI--LDNVIQPLVESIWKRLTLFSGKGHPRAWRG 175
Db 121 SVAGALG-----VGLRRVCMWLSRGQAVPWVEAVRCQSGUGVFSGRDQPAPGEE 169
QY 176 NFDPLWLEHTNEVLEEQ--VSDVEKRRRLMESLRGPADVIRILKSNPALTITAECLKALE 234
Db 170 SFEVLDHTTTEMLHWQGVSERRERRELLSGLRGTAQLVHALLAENPARTAQDCLAALA 229
QY 235 QVFGSVSESSDAQIKFLNTYONPGEKLSAVVIRLEPLLOKVVEKGAIDKDNVNOARLEQV 294
Db 230 QVFGDNESQATIRVKCLTAQQQSGERLSAFVLRLEVLVLLQKAMEKEALASADRVLROM 289
```

```
QY 295 IAGANHSGAIRROLMLTGCAGEGPGPKPLSVAGADP 329
Db 290 LTRAHLTEPLDEAL-----RKLRMAGRSP 313

RESULT 15
US-10-037-860-11
; Sequence 11, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 283
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-11

Query Match 32.6%; Score 564; DB 13; Length 283;
Best Local Similarity 47.0%; Pred. No. 3.2e-48;
Matches 117; Conservative 47; Mismatches 73; Indels 12; Gaps 4;

QY 83 GKGVWVKVLPKPTSDAEFLERLHLFLAREGTVQDVARVLGFQNPPTP---GPMPA 138
Db 3 GKGVWVKVIFKTPNQDTFELERLNLFLKEGQTVSGMFRALGQGVSPATVPFCISPELLA 62
QY 139 EMLNYILDNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPWLHHTNEVLEEQVSDVEK 198
Db 63 HLLGQAMAHAPQPLL-PMRYRKLRFVSGSAVPAPPEESFEVWLQATEIVKEWPTEAEK 121
QY 199 RRLMESLRGPADVIRILKSNPALTITAECLKALEQVSGSVSSRDAQIKFLNTYQNP 258
Db 122 KRWLAESLRGPALDLMHIVQADNPISVEECLEAFKQVFGSLSRRTAQVRYLKYQEEG 181
QY 259 EKLSAYVIRLEPLLOKVVEKGAIDKDNVNOARLEQVIAGANHSGAIRROLML---TGAGE 315
Db 182 EKVSAYVLRLETLRLKAVEKRAIPRRITADQVRLQVWAGA---TLNQMLWCRLRELKDQ 237
QY 316 GPGPKPLSV 324
Db 238 GPPPSFLEL 246
```

Search completed: September 21, 2004, 14:04:06  
Job time : 90.4672 secs

Blank

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 24.461 Seconds  
(without alignments)  
694.369 Million cell updates/sec

Title: US-10-037-860-4

Perfect score: 1729

Sequence: 1 MAMTILLEDWCRGMDVNSORT.....LTGAGEGPGPKPLSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1729	100.0	329	4	US-09-189-527-4
2	747.5	43.2	462	4	US-09-189-527-13
3	441	25.5	195	4	US-09-189-527-7
4	103	6.0	935	4	US-09-914-259-25
5	100.5	5.8	2293	3	US-09-368-590-2
6	97.5	5.6	288	4	US-09-488-039A-12764
7	93.5	5.4	573	4	US-09-328-352-6016
8	92.5	5.3	2431	1	US-07-920-281C-2
9	92.5	5.3	2431	3	US-08-466-277-2
10	91.5	5.3	378	2	US-09-055-097-1
11	91.5	5.3	378	4	US-09-373-902-1
12	91	5.3	565	4	US-09-543-681A-5919
13	90	5.2	688	4	US-09-252-991A-32748
14	89.5	5.2	880	4	US-09-489-039A-12446
15	89	5.1	178	4	US-09-489-039A-11551
16	88.5	5.1	389	4	US-09-252-991A-22086
17	87	5.0	600	3	US-09-212-971-12
18	87	5.0	600	3	US-08-800-929A-12
19	87	5.0	600	4	US-09-617-053A-12
20	87	5.0	1300	4	US-09-543-681A-4501
21	86.5	5.0	373	4	US-09-328-352-7009
22	86	5.0	854	4	US-09-134-000C-4673
23	85	4.9	469	3	US-08-985-335-9
24	85	4.9	469	3	US-09-410-372-9
25	85	4.9	1209	4	US-09-252-991A-25844
26	84.5	4.9	1105	3	US-08-999-774A-2
27	84	4.9	406	4	US-09-328-352-6564

28 84 4.9 871 4 US-09-134-001C-3979 Sequence 3979, Ap  
29 83.5 4.8 332 4 US-09-252-991A-24064 Sequence 24064, A  
30 83.5 4.8 431 4 US-09-543-681A-6055 Sequence 6055, Ap  
31 83 4.8 341 1 US-08-314-309A-19 Sequence 19, Appl  
32 83 4.8 524 3 US-08-557-210A-3 Sequence 3, Appl  
33 83 4.8 539 3 US-08-557-210A-4 Sequence 4, Appl  
34 83 4.8 539 3 US-08-557-210A-5 Sequence 5, Appl  
35 83 4.8 561 4 US-09-489-039A-8807 Sequence 8807, Ap  
36 83 4.8 602 4 US-09-201-936-40 Sequence 40, Appl  
37 83 4.8 602 4 US-09-011-356-40 Sequence 40, Appl  
38 83 4.8 602 4 US-09-672-717-227 Sequence 227, Ap  
39 83 4.8 691 4 US-08-483-941-2 Sequence 2, Appl  
40 83 4.8 691 5 PCT-US91-08442-2 Sequence 2, Appl  
41 83 4.8 735 2 US-08-313-185-48 Sequence 48, Appl  
42 83 4.8 735 2 US-08-459-499-9 Sequence 9, Appl  
43 83 4.8 735 2 US-08-459-499-12 Sequence 12, Appl  
44 83 4.8 735 3 US-09-082-614A-48 Sequence 48, Appl  
45 82.5 4.8 604 2 US-08-511-485-6 Sequence 6, Appl

#### ALIGNMENTS

RESULT 1  
US-09-189-527-4  
; Sequence 4, Application US/09189527A  
; Patent No. 6387639  
; GENERAL INFORMATION:  
; APPLICANT: Jerome B. Posner  
; APPLICANT: Josep O. Dalmau  
; APPLICANT: Myrna R. Rosenfeld  
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma  
; TITLE OF INVENTION: Antibodies  
; FILE REFERENCE: SLK98-01  
; CURRENT APPLICATION NUMBER: US/09/189,527A  
; CURRENT FILING DATE: 1998-11-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 329  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-189-527-4

Query Match 100.0%; Score 1729; DB 4; Length 329;  
Best Local Similarity 100.0%; Pred. No. 1.8e-182;  
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAMTILLEDWCRGMDVNSORTLLVWGIPVNCDBAEETLQAMPOVSYRMILGRMFWREEN 60  
Db 1 MAMTILLEDWCRGMDVNSORTLLVWGIPVNCDBAEETLQAMPOVSYRMILGRMFWREEN 60  
QY 61 AKAAELLELTGAVDYAAIIPREMPGKGVKVLFPKPTSDAEFLERLHLFLAREGWTVDVA 120  
Db 61 AKAAELLELTGAVDYAAIIPREMPGKGVKVLFPKPTSDAEFLERLHLFLAREGWTVDVA 120  
QY 121 RVLGQNPPTPPGPEMPAEMNLNLDNVIQPLVESIWKRLTLFSGKHGPRAWGNFDPW 180  
Db 121 RVLGQNPPTPPGPEMPAEMNLNLDNVIQPLVESIWKRLTLFSGKHGPRAWGNFDPW 180  
QY 181 LEHNEVLEEQVSDVEKRRRLMESLGRPADVIRILKSNPNPAITTAECALKAEQVGSV 240  
Db 181 LEHNEVLEEQVSDVEKRRRLMESLGRPADVIRILKSNPNPAITTAECALKAEQVGSV 240  
QY 241 ESSRDAQIKFLNTYQNFGEKLSAYVIRLEPLQKVEKGAIKDNVQARLEQVIAGNH 300  
Db 241 ESSRDAQIKFLNTYQNFGEKLSAYVIRLEPLQKVEKGAIKDNVQARLEQVIAGNH 300  
QY 301 SGAIRRQLWLTGAGEGPGPKPLSVAGADP 329  
Db 301 SGAIRRQLWLTGAGEGPGPKPLSVAGADP 329

## RESULT 2

US-09-189-527-13  
 ; Sequence 13, Application US/09189527A  
 ; Patent No. 6387639  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jerome B. Posner  
 ; APPLICANT: Josep O. Dalmau  
 ; APPLICANT: Myrna R. Rosenfeld  
 ; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma  
 ; TITLE OF INVENTION: Antibodies  
 ; FILE REFERENCE: SLK98-01  
 ; CURRENT APPLICATION NUMBER: US/09/189,527A  
 ; CURRENT FILING DATE: 1998-11-10  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 13  
 ; TYPE: PRT  
 ; LENGTH: 462  
 ; ORGANISM: homo sapiens  
 ; ORGANISM: homo sapiens  
 US-09-189-527-13

Query Match 43.2%; Score 747.5; DB 4; Length 462;

Best Local Similarity 49.8%; Pred. No. 1.1e-73;

Matches 153; Conservative 49; Mismatches 102; Indels 3; Gaps 2;

QY 7 EDWCRGMDVNSORTLLVWGIPVNCDEAEIETLQAAAPQVS-YRMLGRMFWRRENAKAAL 65  
 DB 1 QDWCRGMDVNSORTLLVWGIPVNCDEAEIETLQAAAPQVS-YRMLGRMFWRRENAKAAL 60  
 QY 66 LEITGADVAAIPREMPGKGGWKVLFKPPSTDAAEFLERLHLFLAREGWTQDVAVLGF 125  
 DB 61 LELAQIDVALLPREIPGKGPWEVIVKPNSDGEFLNRLNLEERRTVSDMNVLS 120  
 QY 126 QNPTPTPGEMPAMLYN--ILDNVIQPLVESIWKRLTLFSGKGHPRAWRGNFDPLWH 183  
 DB 121 DTNCSAPRVTISPEFWTAQTLGAAYQPLLEQMLYRELRFVSGNTTISIFCALAFDAWLEH 180  
 QY 184 TNEVLEEWQVSDVEKERRLMESLRGAADVIRILKSNPAITTAECLEKALEQVFGSVES 243  
 DB 181 TTEMLQMWQVPEGEKRRRLMECLRGALQVSGLRASNASITVECLALQVFGPVESH 240  
 QY 244 RDAQIKFLNTYQNGEKLKSAVIRLEPLQKVEKGAIDKDNVNOARLEQVIAGANHSGA 303  
 DB 241 KIAQVKLCRAYQAGEKVSFVLRLEPLQRAVENVVSRNNVQTLKRVLSGATLPDK 300  
 QY 304 IRRQLWL 310  
 DB 301 LRDKLKL 307

## RESULT 3

US-09-189-527-7  
 ; Sequence 7, Application US/09189527A  
 ; Patent No. 6387639  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jerome B. Posner  
 ; APPLICANT: Josep O. Dalmau  
 ; APPLICANT: Myrna R. Rosenfeld  
 ; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma  
 ; TITLE OF INVENTION: Antibodies  
 ; FILE REFERENCE: SLK98-01  
 ; CURRENT APPLICATION NUMBER: US/09/189,527A  
 ; CURRENT FILING DATE: 1998-11-10  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 7  
 ; TYPE: PRT  
 ; LENGTH: 195  
 ; ORGANISM: homo sapiens  
 ; ORGANISM: homo sapiens  
 US-09-189-527-7

Query Match 25.5%; Score 441; DB 4; Length 195;

Best Local Similarity 46.4%; Pred. No. 2.1e-40;

Matches 90; Conservative 37; Mismatches 61; Indels 6; Gaps 4;  
 QY 3 MTLLEDWCRGMDVNSORTLLVWGIPVNCDEAEIETLQAAAPQVS-YRMLGRMFWRRENA 61  
 DB 2 LALLEDWCRIMSDQKSLMTGIPADFEAEIQEVLQETLSLGRYLLGKIFRQENA 61  
 QY 62 KAALELTGADVAAIPREMPGKGGWKVLFKPPSTDAAEFLERLHLFLAREGWTQDVAV 121  
 DB 62 NAVLELLEDTDVSAIPSEVQKGGWKVIFKTPNQDTFELRLNLFLEKGGQTVSGMFR 121  
 QY 122 VLGFQ--NPTPTP--GPMPAEMLYILDNVIQPLVESIWKRLTLFSGKGHPRAWRGNF 177  
 DB 122 ALGQEAISPATVPCISPELLAHLQAAHAPOPLL-PMYRKRLRVFSGSAVPAPEESF 180  
 QY 178 DPWLEHTNEVLEEM 191  
 DB 181 EVMLEQATEIVKEM 194

## RESULT 4

US-09-914-259-25  
 ; Sequence 25, Application US/09914259  
 ; Patent No. 6495336  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Makowski, Lee  
 ; APPLICANT: Hyman, Paul  
 ; APPLICANT: Williams, Mark  
 ; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES  
 ; FILE REFERENCE: 8471-010-999  
 ; CURRENT APPLICATION NUMBER: US/09/914,259  
 ; CURRENT FILING DATE: 2000-11-21  
 ; NUMBER OF SEQ ID NOS: 180  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 25  
 ; LENGTH: 935  
 ; TYPE: PRT  
 ; ORGANISM: Syncephalastrum racemosum  
 ; ORGANISM: Syncephalastrum racemosum  
 US-09-914-259-25

Query Match 6.0%; Score 103; DB 4; Length 935;

Best Local Similarity 21.7%; Pred. No. 0.065;

Matches 69; Conservative 56; Mismatches 121; Indels 72; Gaps 15;  
 QY 16 NSORTLLVWGIPVNCDEAEIETLQAAAPQVS-YRMLGRMFWRRENAKAALEL----- 68  
 DB 297 NSRTTIIINCSPPSSYNEAETLSTLRFGARAKSIKNKAKV-----NADLSPAELKALIKV 351  
 QY 69 -TGADVAAIPREMPGKGGWKVLFKPPSTDAAEFLERLHLFLAREG-WTVQDVAVRLGFQ 126  
 DB 352 KSEAVTYQTYIAALEGEVNVWRTGGTVP-----EGKWVTMDKVKSGDFA 395  
 QY 127 NPTPTPGMPAEMLYILDNVIQPLVESIWKRLTLFSGKGHPRAWRGNFDPLWLEHTNE 186  
 DB 396 GLPPAPGFKSP-----VSDEGSRPATPV-----PTLEKDREREFIKENE 435  
 QY 187 VLEEWQVSDVE-----KRRRLMESLR--GPAADVIRILKSNPAIT--AECLKALEQVF 237  
 DB 436 LMD--QISEKETELTNREKLLESIREMGYYKEQSVTKENQOQMTSELRLQLQKV-- 492  
 QY 238 GSVESSRDAQIKFLNTYQNGEKLKSAVIRLEPLQKV--VEKGAIDKNVNO--ARLEQ 293  
 DB 493 -SYESKENAIT--VDSLKEANQDLMALEELKKNLSEMROAHKDATDSDKEKRAEKAAQ 549  
 QY 294 VIAGANHSGAI---RRQL 308  
 DB 550 MMSGFDPGSLINDKERQI 567

## RESULT 5

US-09-368-590-2  
 ; Sequence 2, Application US/09368590  
 ; Patent No. 6187563  
 ; GENERAL INFORMATION:



```

; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/920,281C
; FILING DATE: 13-AUG-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 828-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-920-281C-2

```

```

Query Match      5.3%; Score 92.5; DB 1; Length 2431;
Best Local Similarity 19.6%; Pred. No. 4.6;
Matches 74; Conservative 50; Mismatches 112; Indels 141; Gaps 15;

QY 52 GRMFWREENAKAALLBELTGAVDYAAIPREMPGGGKGVKVLFKPPTSDAEFLERLHLFL-- 109
Db 865 GKMRITNPNCKPIIIDTTG-----QTKPKPGDIVLTCFRGWAKQLQLDYRGHEVMTA 916
QY 110 -AREGTVQDVAVRLGFQNPPTPGPEMPAEMLYILDNVIQPLVESIWYKRLTLFSGKG 168
Db 917 ASQGLTRKGVAVRQKVNENPLYAP--ASEHVNVLTRTEDRLV-----WKTILA----- 964
QY 169 HPRAMRGNDPWLHETNEV-----LEEWQVSDVEKRRRLMESLRGPAADV----- 213
Db 965 -----GDPWKVLSNIPQGNFTATLEWQ-----EEHDKIMKVIEGPAAPVDAFONKA 1012
QY 214 -----IRILKSNPAITTA-----ECLKALEQV----- 236
Db 1013 NVCWAKSLVPDLTAGIRLTAEWSTIIITAFKEDRAYSPVVALNEICTKYGYVDLDSGLF 1072
QY 237 -----FGSVESRR-DAQIKFLNTYONPEKLSAVVIRLEPL 271
Db 1073 SAPKVSLEYENNHNDRPGRMVGFNAATAARLEAHTFLKGQWHTGKQAVIAERKIQPL 1132
QY 272 --LQKVE-----KGAIDKDNVNOARLEQVIAGANHSAGI--RRQLWLT 311
Db 1133 SVLDNVIPINRRLPHALVAEYKTVKGRVEMLVNKGVRGYHLLVSEYNLALPRRRVTWLS 1192
QY 312 GAGEGPGPKPLSVAGAD 328
Db 1193 -----PLNVTGAD 1200

```

```

RESULT 9
US-08-466-277-2
; Sequence 2, Application US/08466277
; Patent No. 6190666
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik
;             Liljestrom, Peter
; TITLE OF INVENTION: DNA Expression Systems Based on
;                   Alphaviruses
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,277
; FILING DATE: 06-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/920,281
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 828-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-466-277-2

Query Match      5.3%; Score 92.5; DB 3; Length 2431;
Best Local Similarity 19.6%; Pred. No. 4.6;
Matches 74; Conservative 50; Mismatches 112; Indels 141; Gaps 15;

QY 52 GRMFWREENAKAALLBELTGAVDYAAIPREMPGGGKGVKVLFKPPTSDAEFLERLHLFL-- 109
Db 865 GKMRITNPNCKPIIIDTTG-----QTKPKPGDIVLTCFRGWAKQLQLDYRGHEVMTA 916
QY 110 -AREGTVQDVAVRLGFQNPPTPGPEMPAEMLYILDNVIQPLVESIWYKRLTLFSGKG 168
Db 917 ASQGLTRKGVAVRQKVNENPLYAP--ASEHVNVLTRTEDRLV-----WKTILA----- 964
QY 169 HPRAMRGNDPWLHETNEV-----LEEWQVSDVEKRRRLMESLRGPAADV----- 213
Db 965 -----GDPWKVLSNIPQGNFTATLEWQ-----EEHDKIMKVIEGPAAPVDAFONKA 1012
QY 214 -----IRILKSNPAITTA-----ECLKALEQV----- 236
Db 1013 NVCWAKSLVPDLTAGIRLTAEWSTIIITAFKEDRAYSPVVALNEICTKYGYVDLDSGLF 1072
QY 237 -----FGSVESRR-DAQIKFLNTYONPEKLSAVVIRLEPL 271
Db 1073 SAPKVSLEYENNHNDRPGRMVGFNAATAARLEAHTFLKGQWHTGKQAVIAERKIQPL 1132
QY 272 --LQKVE-----KGAIDKDNVNOARLEQVIAGANHSAGI--RRQLWLT 311
Db 1133 SVLDNVIPINRRLPHALVAEYKTVKGRVEMLVNKGVRGYHLLVSEYNLALPRRRVTWLS 1192
QY 312 GAGEGPGPKPLSVAGAD 328
Db 1193 -----PLNVTGAD 1200

RESULT 10
US-09-055-097-1
; Sequence 1, Application US/09055097
; Patent No. 5955282
; GENERAL INFORMATION:

```

```

; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,097
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0490 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: DUODNOT02
; CLONE: 1705085
; US-09-055-097-1

```

```

Query Match          5.3%; Score 91.5; DB 2; Length 378;
Best Local Similarity 23.9%; Pred. No. 0.27;
Matches 61; Conservative 26; Mismatches 89; Indels 79; Gaps 15;

Qy 11 RGMVY-----NSQRTLLVW---GIPVNCDEAEIETLQAAMPQVSYR-----MLGR 53
Db 101 RGLRVQTLFLGEPNACHP--VWGSQSDLASESAQGDILQAAP-QDSYRNLTSLTSG 157
Qy 54 MFWREENAKAALLELTGAVD-YAALPREMP-----GKGWVKVLFKPP-TSDAEFLERL 105
Db 158 LNWAKEHCMPARYVLKTDVVVNVNVPVSELVLRGGWGWQWSTEREQREAEQGGQVL 217
Qy 106 H-----LFLAREGWTQDVVARVLGFQNPPTPGPEMPAEMLYINDNVIQPLVESIWK 159
Db 218 HSEVPVLLYLRGVHVRV-----NPSRTPGGR-----HRVSEEQW-- 251
Qy 160 RLTLFSGKHPRAMGNFDPMLHETNEVLEEQVS---DVEKRRRLMESLRGPADVIRI 216
Db 252 -----PHTW-GPPFPYASGTGYLSASAVQLILKVASRAPLL-----PLEDVFGV 295
Qy 217 LKSNNPALTITAECLK 231
Db 296 VSARRGGLAPTQCVK 310

```

```

RESULT 11
US-09-373-902-1
; Sequence 1, Application US/09373902
; Patent No. 6649737

```

```

; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Guegler, Karl J.
; Corley, Neil C.
; Shah, Purvi
; Patterson, Chandra
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/373,902
; FILING DATE: 12-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,097
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0490 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: DUODNOT02
; CLONE: 1705085
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-09-373-902-1

```

```

Query Match          5.3%; Score 91.5; DB 4; Length 378;
Best Local Similarity 23.9%; Pred. No. 0.27;
Matches 61; Conservative 26; Mismatches 89; Indels 79; Gaps 15;

Qy 11 RGMVY-----NSQRTLLVW---GIPVNCDEAEIETLQAAMPQVSYR-----MLGR 53
Db 101 RGLRVQTLFLGEPNACHP--VWGSQSDLASESAQGDILQAAP-QDSYRNLTSLTSG 157
Qy 54 MFWREENAKAALLELTGAVD-YAALPREMP-----GKGWVKVLFKPP-TSDAEFLERL 105
Db 158 LNWAKEHCMPARYVLKTDVVVNVNVPVSELVLRGGWGWQWSTEREQREAEQGGQVL 217
Qy 106 H-----LFLAREGWTQDVVARVLGFQNPPTPGPEMPAEMLYINDNVIQPLVESIWK 159
Db 218 HSEVPVLLYLRGVHVRV-----NPSRTPGGR-----HRVSEEQW-- 251
Qy 160 RLTLFSGKHPRAMGNFDPMLHETNEVLEEQVS---DVEKRRRLMESLRGPADVIRI 216
Db 252 -----PHTW-GPPFPYASGTGYLSASAVQLILKVASRAPLL-----PLEDVFGV 295
Qy 217 LKSNNPALTITAECLK 231
Db 296 VSARRGGLAPTQCVK 310

```

```

RESULT 12
US-09-543-681A-5919

```

```

; Sequence 5919, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5919
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-5919

```

Query Match	5.3%;	Score 91;	DB 4;	Length 565;
Best Local Similarity	22.7%;	Pred. No. 0.6;		
Matches	56;	Conservative	33;	Mismatches 76; Indels 82; Gaps 12;
QY	90	VLEPKPTS--DABFELRLHLFLAREGWTVQDVARVLGFQNPPTTPGPEMPAEMLYILDN	147	
	:	:::	:	:::
Db	194	LLLLDEPNHLDAAESVAMLERFLHDYEGTVAITHD-----RYFLDN	234	
QY	148	VIOPLESYIWKLLTLFSGKHCPRAWCGNFDPMLEHTNEVLIEWOVSDVEKRRRLMESLR	207	
	:	:::	:	:::
Db	235	VAGWILE-----LDRGSGIP--WEGNTSSMLDEQKDARLEQAATAEARHKSIQKEL-	283	
QY	208	GPAADVIRILKSNPPAITTAECCLKALREQVFSGVESSEDAQIKPLNT--YQN-----	256	
	:	:::	:	:::
Db	284	----EWIR---QNPFKGRQAK-----GKARLARFEELNSVDYQKRNETSELF	323	
QY	257	-PGEKLSAYVIRLEPL-----LQKVVEKGA-----IDKNVNVOARLEOVIVAGAN	299	
	:	:::	:	:::
Db	324	PPPERUGDKVLEVSHLTKSVGDVRVLIIDDLFSFPKGAIUVGLIGPGAGACKSTLFMWISQGE	383	
QY	300	H--SGAI	304	
	:	:	:	
Db	384	QPDSGTV	390	

```

RESULT 13
US-09-252-991A-32748
; Sequence 32748, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCES: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32748
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32748

```

```

Query Match      5.23; Score 90; DB 4; Length 688;
Best Local Similarity 25.23; Pred. No. 1.1;
Matches 52; Conservative 17; Mismatches 59; Indels 78; Gaps 14;

Qy 3 MTLLEDRCGMNDVNSQRTLLVWGIPVNCDEAEIETIIQAAMPQVSYEMLGKMFV----- 56
Db 535 MESPEQNMAMDT---RTYIADDDILAKVDRAAMATSLSTRVPLIDHRVV--ELAKWKMPIL 590

```

```

Qy 57 --REENKAALLELTGADVAAIIPREM---PGKG---GWKVLFPKPTSDAEFL---E 103
Db 591 KIENNCGKMLLRQ-----VLVQHVPRELIERPKWGFAPLDTW--LRGLREWAESLLAE 644
Qy 104 RLHLFLAREGWTQDVARVLGFQNPPTTPGPEMPAEMNLVDNVPGLVESIHWKRLTL 163
Db 645 RLR--RE-----GYLRPEP-----IRRAWOEHL-- 665
Qy 164 FSKGHPRAWRG-----NFDPWLEH 183
Db 666 -SGK---RNWOGPLWNVIMFQAWLEH 687

```

RESULT 14  
US-09-489-039A-12446  
; Sequence 12446, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12446  
; LENGTH: 880  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12446

Query Match	5.2%; Score 89.5; DB 4; Length 880;
Best Local Similarity	20.9%; Pred.No.1.8;
Matches	63; Conservative 46; Mismatches 126; Indels 67; Gaps 15;

  

Qy	49	RMLGRMPWRENAKAALLEL-----TGAVDYAAIPREMPGKGWGVKVLFPKPTSDAEFL	102
Db	1	RYVGRIMTWARRTVLVEEYKRVHVAERAAAGIVAKPLDATQMAALVELLKNPPAGEEFL	60
Qy	103	ERLHLFLAREGWTVDVARV-LGF-----QNPPTFGPEMPAEMLNLYILDNV-IQPLV	153
Db	61	--LDLLINRVPPGVDEAAVYKAGFLAAIAGEATSPLVTPPEKAVELLGTMQGGINHPLI	118
Qy	154	ESTIWKRLTLFSGKG-----HPRWRGNFDPWLEHTNEVLEWQVSDVE--	197
Db	119	DALDDAKLAPIAAKALSHLTLLMFDFYDVEBEKAKAGN-----EHAQVMQSW--ADA	171
Qy	198	-KRRRLMESLRGPAADVIRIL-KSNNPATITAECLKALEQVFGVESRDAQIKFLNTYQ	255
Db	172	LNRPLQAEKI---TWTVKVTGETNTDGLSPAP-----DAWGRPDIPLHALAMLK	218
Qy	256	NPEKLSAYVIRLEP-----LLQKVVEKGATDKNNVNOARLEQVI-AGANHSGAIRRQLWL	310
Db	219	NPRE-----GTEPDQGVVGPIKQIEALQOKQYPLAYGVVGTGSSRSKATNSVLWF	271
Qy	311	TG 312	
Db	272	MG 273	

```

RESULT 15
US-09-489-039A-11551
; Sequence 11551, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gaty Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27

```



Blank

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 : Search time 60.7447 Seconds  
(without alignments)  
1530.308 Million cell updates/sec

Title: US-10-037-860-4

Perfect score: 1729

Sequence: 1 MAMTLEDWCRGMDVNSQRT.....LTGAGEGFGPKPLSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1704	98.6	329	3 AAB12525	Aab12525 Human Ma1
2	1602	92.7	353	4 AAB74701	Aab74701 Human mem
3	887.5	51.3	351	4 AAE01340	Aae01340 Human gen
4	887.5	51.3	351	4 AAB74695	Aab74695 Human mem
5	887.5	51.3	351	4 AAU08664	Aau08664 Human NOV
6	874.5	50.6	312	3 AAB43023	Aab43023 Human ORF
7	770.5	44.6	364	7 ADC08977	Adc08977 Oncogene
8	766.5	44.3	455	5 ABB05727	Abb05727 Human sig
9	766.5	44.3	463	3 AAB12529	Aab12529 Human Ma5
10	766.5	44.3	463	3 AAB42315	Aab42315 Human ORF
11	744	43.0	452	6 AAO16179	Aao16179 Human pro
12	620	35.9	399	6 ADA54410	Ada54410 Human pro
13	620	35.9	399	6 ABG99947	Abg99947 Human nov
14	570.5	33.0	280	4 AAE01336	Aae01336 Human gen
15	565	32.7	283	3 AAB12528	Aab12528 Human Ma4
16	540	31.2	237	4 AAB94854	Aab94854 Human Ma4
17	441	25.5	195	3 AAB12526	Aab12526 Human Ma2
18	399	23.1	110	4 AAO01787	Aao01787 Human pol
19	391.5	22.6	439	6 ABP75736	Abp75736 Human sec
20	335.5	19.4	403	5 ABG97495	Abg97495 Human NOV
21	335.5	19.4	403	5 AAM51624	Aam51624 KIA0883-
22	335.5	19.4	403	6 ABO14772	Abo14772 Novel hum
23	335	19.4	403	6 ABO14773	Abo14773 Novel hum
24	332	19.2	402	4 AAB60478	Aab60478 Human cel
25	326	18.9	337	4 AAM25693	Aam25693 Human pro

## ALIGNMENTS

## RESULT 1

AAB12525  
ID AAB12525 standard; protein; 329 AA.

XX AAB12525;

DT 02-NOV-2000 (first entry)

XX Human Ma1 protein SEQ ID NO:4.

XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;  
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;  
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;  
KW germ-cell tumour.

OS Homo sapiens.

PN JP2000146982-A.

XX 26-MAY-2000.

XX 10-NOV-1999; 99JP-00320171.

XX 10-NOV-1998; 98US-00189527.

XX (SLOK ) SLOAN KETTERING INST CANCER RES.

XX WPI; 2000-468119/41.

XX N-PSDB; AAA60833.

PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic  
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test  
PT sample for the presence or absence of antibodies to a Ma family  
polypeptide.

XX Claim 48; Fig 1; 27pp; Japanese.

CC The present invention describes a method for diagnosing a paraneoplastic  
CC syndrome or neoplasm. The method comprises assessing a test sample for  
CC the presence or absence of antibodies to a Ma family polypeptide (I). The  
CC method is used to diagnose a paraneoplastic syndrome especially  
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or  
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing  
CC antibodies to (I) preferably Ma1, which is indicative presence of breast  
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular  
CC cancer and germ-cell tumours or Ma2, which is indicative of testicular  
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the  
CC Ma1 protein as given in the present invention

Aab12527 Human Ma3  
Aam14937 Peptide #  
Abb33906 Peptide #  
Aam27367 Peptide #  
Abb28727 Peptide #  
Abb19347 Protein #  
Aam67075 Human bon  
Aam54673 Human bra  
Abg48741 Human liv  
Aam02665 Peptide #  
Abg36734 Human pep  
Abg19651 Novel hum  
Abu14259 Novel hum  
Abu52641 Human bra  
Aad31124 Human nov  
Aau28186 Novel hum  
Abu21717 Protein e  
Aay75199 Neisseria  
Abu37227 Protein e

26 256.5 14.8 149 3 AAB12527  
27 246.5 14.3 116 4 AAM14937  
28 246.5 14.3 116 4 ABB33906  
29 246.5 14.3 116 4 AAM27367  
30 246.5 14.3 116 4 ABB28727  
31 246.5 14.3 116 4 ABB19347  
32 246.5 14.3 116 4 AAM67075  
33 246.5 14.3 116 4 AAM54673  
34 246.5 14.3 116 4 AAG48741  
35 246.5 14.3 116 4 AAM02665  
36 246.5 14.3 116 5 ABG36734  
37 221 12.8 615 4 ABG19651  
38 221 12.8 615 4 ABG14259  
39 192 11.1 538 4 ABU52641  
40 192 11.1 538 7 ADC31124  
41 133 7.7 584 4 AAU28186  
42 115.5 6.7 555 6 ABU21717  
43 114.5 6.6 555 6 ABU22846  
44 114 6.6 373 3 AAY75199  
45 114 6.6 558 6 ABU37227

```
XX SQ Sequence 329 AA;
Query Match 98.6%; Score 1704; DB 3; Length 329;
Best Local Similarity 98.8%; Pred. No. 1.4e-165;
Matches 325; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMLGRMPWREN 60
DB 1 MAMTLLDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMLGRMPWREN 60
QY 61 AKAAELLELTGAVDYAAIPREMPGKGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
DB 61 AKAAELLELTGAVDYAAIPREMPGKGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
QY 121 RVLGFGNPTPTGPEMPAEMLNLYLDNVQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180
DB 121 RVLGFGNPTPTGPEMPAEMLNLYLDNVQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180
QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATITTAECLEQVFGSV 240
DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATITTAECLEQVFGSV 240
QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
QY 301 SGAIRQLWLTGAGEGPGKPLSVAGADP 329
DB 301 SGAIRQLWLTGAGEGPGKPLSVAGADP 329

RESULT 2
AAB74701
ID AAB74701 standard; protein; 353 AA.
AC AAB74701;
XX
XX 12-JUN-2001 (first entry)
DE
DE Human membrane associated protein MEMAP-7.
XX
XX Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
XX antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;
XX antiarteriosclerotic; gene therapy; cell proliferative disorder;
XX autoimmune disorder; inflammatory disorder; neurological disorder;
XX gastrointestinal disorder; cancer; inflammation; atherosclerosis;
XX epilepsy; diarrhoea.
XX
XX Homo sapiens.
OS
XX WO200112662-A2.
PN
XX 22-FEB-2001.
PD
XX
XX 14-AUG-2000; 2000WO-US022315.
PF
XX
XX 17-AUG-1999; 99US-0149641P.
PR
XX 09-NOV-1999; 99US-0164203P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;
XX Baughn MR, Lu DAM, Patterson C;
XX
XX WPI; 2001-168860/17.
DR
XX N-PSDB; AAF81747.
XX
XX Isolated polypeptide with a human membrane associated protein sequence is
XX useful for the diagnosis, prevention and treatment of cell proliferative,
XX autoimmune/inflammatory, neurological and gastrointestinal disorders.
XX
XX Claim 1; Page 119-120; 173pp; English.
XX PS
```

```
XX AAF81741 to AAF81777 encode the human membrane associated proteins
CC (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,
CC antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and
CC antiarteriosclerotic activities, which can be used in gene therapy.
CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
CC associated with decreased expression of functional MEMAP and antagonists
CC of MEMAP are used to treat a disease or condition associated with
CC overexpression of functional MEMAP. These disorders include cell
CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal
CC disorders. The MEMAP polynucleotides and proteins are also used for the
CC diagnosis of these disorders. Specific examples of these disorders
CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
CC MEMAP proteins can be used to screen for compounds which specifically
CC bind MEMAP including antibodies, oligonucleotides, proteins and small
CC molecules. MEMAP polynucleotides can be used to prepare transgenic
CC animals which can be studied to provide information concerning human
CC disease. Anti-MEMAP antibodies are useful in immunoassays for the
CC detection of MEMAP protein and can be used as antagonists to treat or
CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
CC can be delivered to target cells with genetic abnormalities with respect
CC to the expression of MEMAP to treat or prevent a disorder associated with
CC MEMAP
XX
XX SQ Sequence 353 AA;
Query Match 92.7%; Score 1602; DB 4; Length 353;
Best Local Similarity 96.6%; Pred. No. 4.4e-155;
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMLGRMPWREN 60
DB 1 MAMTLLDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMLGRMPWREN 60
QY 61 AKAAELLELTGAVDYAAIPREMPGKGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
DB 61 AKAAELLELTGAVDYAAIPREMPGKGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
QY 121 RVLGFGNPTPTGPEMPAEMLNLYLDNVQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180
DB 121 RVLGFGNPTPTGPEMPAEMLNLYLDNVQPLVESIWKYKRLTLFSGRDIPGGEETFDPW 180
QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATITTAECLEQVFGSV 240
DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATITTAECLEQVFGSV 240
QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
QY 301 SGAIRQLWLTGAGEGPGP 319
DB 301 SGAIRQLWLTGAGEGPA 319

RESULT 3
AAE01340
ID AAE01340 standard; protein; 351 AA.
XX
XX AAE01340;
XX
XX 17-JUL-2001 (first entry)
DT
XX Human gene 22 encoded secreted protein fragment, SEQ ID NO:205.
DE
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiogenic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX
```

endocrine disorder; infection; wound healing; vulnery; cell culture;  
chemotaxis; food additive; gene therapy; binding partner identification.  
Homo sapiens.  
WO200134769-A2.  
17-MAY-2001.  
01-NOV-2000; 2000WO-US030040.  
05-NOV-1999; 99US-0163580P.  
30-JUN-2000; 2000US-0215130P.  
(HUMA-) HUMAN GENOME SCI INC.  
Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;  
WPI; 2001-308781/32.  
New isolated nucleic acid molecule encoding a human secreted protein is  
used in preventing, treating or ameliorating a medical condition.  
Disclosure; Page 46; 519pp; English.  
AAD05121-AAD05203 represent cDNAs corresponding to 24 human secreted  
protein genes, and AAE01232-AAE01311 represent the proteins they encode.  
AAE01312-AAE01340 represent human secreted protein variants or fragments.  
The secreted proteins and their genes are useful for preventing, treating  
or ameliorating medical conditions, e.g., by protein or gene therapy.  
Pathological conditions can be diagnosed by determining the amount of the  
new protein in a sample or by determining the presence of mutations in  
the new genes. Specific uses are described for each of the 24 genes,  
based on the tissues in which they are most highly expressed, and include  
developing products for the diagnosis or treatment of proliferative  
disorders, cancer, tumours, foetal and developmental abnormalities,  
haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
angiogenic disorders, kidney disorders, gastrointestinal disorders,  
pregnancy-related disorders, endocrine disorders, and infections. The  
proteins can also be used to aid wound healing and epithelial cell  
proliferation, to prevent skin aging due to sunburn, to maintain organs  
before transplantation, for supporting cell culture of primary tissues,  
to regenerate tissues, to identify their cognate ligands or binding  
partners, and in chemotaxis, and can be used as a food additive or  
preservative to modify storage properties. Antibodies specific for a  
protein of the invention can be used in alleviating symptoms associated  
with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The  
present sequence represents a human secreted protein fragment referred to  
in the disclosure of the invention  
Sequence 351 AA;  
Query Match  
Best Local Similarity 51.3%; Score 887.5; DB 4; Length 351;  
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;  
QY 1 MAMTLLDWCRCGMDVNSORTLLVGMIPVNCDAEIEETLQAM-PQVSRYMLGRMFREE 59  
DB 1 MTLRLLEDWCRCGMDVNSORTLLVGMIPVNCDAEIEETLQAM-PQVSRYMLGRMFREE 60  
QY 60 NAKAALLETGAVDVAALPREMPGKGVKWLFPKPTSDAFLERLHLFLAREGTVQDV 119  
DB 61 NRKVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPDNTFLSLNEFLAGEGTVGEL 120  
QY 120 ARVLGFQNP--PTPG--PMPAEMVILNDVIOPLVESWYKRLTLFSGKHGPRW 175  
DB 121 SRALGHENGSLDPEQGMIPMPWAPMLAQALE-ALQPALQCLKYKKLRFVSGRESPEEGEE 179

QY 176 NFDPMLEHTNEVLEBQWSDVEKRRRLMESLRGPAADVIRILKSNPNPAITAECLKALRQ 235  
DB 180 EFGRMWFHTTQMIKAWQVDPDEKRRRLLESRLGPAADVIRILKSNPNPAITAECLKALRQ 239  
QY 236 VFGSVESRDAQIKFLNTYQNFGEKLSAYVIRLEPLQKVKVEGAIDKDNVQARLEQVI 295  
DB 240 VFGVTDNPRELQVKYLTQKDEEKLKSAVLEPLQKVKVEGAIDKDNVQARLEQVI 299  
QY 296 AGANHSIGAIRROLWLTGAGEGPG 318  
DB 300 AGAVHK-TIRRELNPEDGPAFG 321  
RESULT 4  
AAB74695  
ID AAB74695 standard; protein; 351 AA.  
AC AAB74695;  
XX  
DT 12-JUN-2001 (first entry)  
DE Human membrane associated protein MEMAP-1.  
XX  
KW Human; membrane associated protein; MEMAP; diagnosis; cytostatic;  
antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;  
KW antiarteriosclerotic; gene therapy; cell proliferative disorder;  
KW autoimmune disorder; inflammatory disorder; neurological disorder;  
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;  
KW epilepsy; diarrhoea.  
OS Homo sapiens.  
XX  
PN WO200112662-A2.  
XX  
PD 22-FEB-2001.  
XX  
PF 14-AUG-2000; 2000WO-US022315.  
PR 17-AUG-1999; 99US-0149641P.  
PR 09-NOV-1999; 99US-0164203P.  
XX (INCY-) INCYTE GENOMICS INC.  
PI Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;  
PI Baughn MR, Lu DM, Patterson C;  
XX  
DR WPI; 2001-168860/17.  
DR N-PSDB; AAF81741.  
XX Isolated polypeptide with a human membrane associated protein sequence is  
useful for the diagnosis, prevention and treatment of cell proliferative,  
PT autoimmune/inflammatory, neurological and gastrointestinal disorders.  
XX  
PS Claim 1; Page 114-115; 173pp; English.  
XX  
CC AAF81741 to AAF81777 encode the human membrane associated proteins  
(MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,  
CC antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and  
CC antiarteriosclerotic activities, which can be used in gene therapy.  
CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition  
CC associated with decreased expression of functional MEMAP and antagonists  
CC of MEMAP are used to treat a disease or condition associated with  
CC overexpression of functional MEMAP. These disorders include cell  
CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal  
CC disorders. The MEMAP polynucleotides and proteins are also used for the  
CC diagnosis of these disorders. Specific examples of these disorders  
CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.  
CC MEMAP proteins can be used to screen for compounds which specifically  
CC bind MEMAP including antibodies, oligonucleotides, proteins and small  
CC molecules. MEMAP polynucleotides can be used to prepare transgenic  
CC animals which can be studied to provide information concerning human  
CC disease. Anti-MEMAP antibodies are useful in immunoassays for the  
CC detection of MEMAP protein and can be used as antagonists to treat or

CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP  
CC can be delivered to target cells with genetic abnormalities with respect  
CC to the expression of MEMAP to treat or prevent a disorder associated with  
CC MEMAP

XX Sequence 351 AA;

Query Match 51.3%; Score 887.5; DB 4; Length 351;  
Best Local Similarity 55.4%; Pred. No. 7.7e-82;  
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;

QY 1 MAMTLLDRCRGMDVNSQRTLLVWGPVNCDEAIEBETLQAAM-PQVSVMGLGRMFWRREE 59  
DB 1 MTLRLLEDWCRGMDVNSQRTLLVWGPVNCDEAIEBETLQAAM-PQVSVMGLGRMFWRREE 60  
QY 60 NAKAALLELTGAVDYAAIPREMPKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDV 119  
DB 61 NRKVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSRLNEFLAGEGWTGEL 120  
QY 120 ARVLGFQNPPT--PTPG--PEMPAEMLNLYILDVNIQPLVESIWKRLTLFSGKHPRAWRG 175  
DB 121 SRAIGHENGSLDPEQGMIPEMWAPMLAQALE-ALQPALQCLKYKLVFSGRSPFEGEE 179  
QY 176 NFDPMLEHTNEVLEEQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 235  
DB 180 EFGRWMFHTTQMIKAMQVDPVDEKRRRLLESRLRGPAALDVIRLVKINNPLITVDCEQLALEE 239  
QY 236 VFGSVSSRDQAQKFLNTYQNPCKLSAVVIRLEPLLOKVVEKGAIDKONVQARLEQVI 295  
DB 240 VFGVTDNPRELVQKYLTTTQKDEKLSAVVIRLEPLLOKVVEKGAIDKONVQARLEQVI 299  
QY 296 AGANHSGAIRRQLWLTGAGEGPG 318  
DB 300 AGAVHK-TIRRELNPEDGPAFG 321

RESULT 5  
AAU08664  
ID AAU08664 standard; protein; 351 AA.  
XX AAU08664;  
AC AAU08664;  
XX 18-DEC-2001 (first entry)  
XX Human NOV8 protein.  
XX Human; NOV8; cytostatic; nontropic; neuroprotective; vulnary;  
KW cerebroprotective; antiparkinsonian; hypotensive; antidiabetic;  
KW antidiabetic; antipsoriatic; antiinflammatory; immunosuppressive;  
KW antithrombotic; dermatological; cancer; neurological disorder;  
KW Parkinson's disease; diabetes mellitus; asthma; enamel defect;  
KW immune disorder; autoimmune disease; respiratory disorder; bone disorder;  
KW musculoskeletal disorders; leukaemia; lymphoma;  
KW cell growth regulation disorder; lesional psoriatic skin;  
KW atherosclerosis; abdominal aortic aneurysm.

OS Homo sapiens.  
XX WC200168851-A2.  
XX 20-SEP-2001.  
XX 12-MAR-2001; 2001WO-US007735.  
XX 10-MAR-2000; 2000US-0188277P.  
XX 10-MAR-2000; 2000US-0188316P.  
XX 14-MAR-2000; 2000US-0189139P.  
XX 14-MAR-2000; 2000US-0189140P.  
XX 17-MAR-2000; 2000US-0190231P.  
XX 17-MAR-2000; 2000US-0190401P.  
XX (CURA-) CURAGEN CORP.

PI Padigaru M, Vernet CAM, Fernandes E, Shimkets RA, Spaderma SK;  
PI Majumder K, Li L;  
XX WPI; 2001-570869/64.  
DR N-PSDB; AAS13342.

XX Novel polypeptides and nucleic acids homologous to members of collagen,  
PT potassium channel, tuftelin family of proteins for diagnosing, treating  
PT cancer, atherosclerosis, neurological, skin and enamel defect disorders.  
XX Claim 1; Page 29; 128pp; English.

XX The invention relates to isolated NOVX (NOVX1-11) polypeptides and the  
CC polynucleotides that encode them. NOVX polypeptides, polynucleotides and  
CC anti-NOVX antibodies are useful for treating or preventing a pathology  
CC associated with NOVX polypeptide in humans and for treating a syndrome  
CC associated with human disease e.g. disorders characterised by altered  
CC cell motility, proliferation and migration e.g. cancer, angiogenesis and  
CC wound healing (NOV1-3), neurological disorders, e.g. episodic ataxia,  
CC autosomal dominant myokymia, stroke, Parkinson's disease, Alzheimer's  
CC disease, non-insulin dependent diabetes mellitus, asthma, hypertension  
CC and seizure (NOV4), enamel defects, such as amelogenesis imperfecta and  
CC disorders involving enamel defects, including hypoplasia and  
CC hypomineralisation (NOV5-7), paraneoplastic neurological disorders, e.g.  
CC paraneoplastic limbic of brain-stem encephalitis occurring during  
CC testicular cancer, diabetes, reproductive health, metabolic and endocrine  
CC disorders, gastrointestinal disorders, immune disorders and autoimmune  
CC diseases, respiratory disorders, bone disorders, musculoskeletal  
CC disorders, leukaemia/lymphoma and tissue/cell growth regulation disorders  
CC (NOV8), lesional psoriatic skin (NOV9-10) and atherosclerosis, abdominal  
CC aortic aneurysm and neurological disorders (NOV11). NOVX polypeptide is  
CC also useful for identifying an agent that binds to it and a cell  
CC expressing NOVX polypeptide is useful for identifying a therapeutic agent  
CC for use in treatment of a NOVX related pathology. The antibodies and a  
CC polypeptide having 95% sequence identity to NOVX polypeptide are useful  
CC for treating a pathological state in a mammal. The present sequence  
CC represents NOV8, a possible neuronal antigen-like protein  
XX Sequence 351 AA;

Query Match 51.3%; Score 887.5; DB 4; Length 351;  
Best Local Similarity 55.4%; Pred. No. 7.7e-82;  
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;

QY 1 MAMTLLDRCRGMDVNSQRTLLVWGPVNCDEAIEBETLQAAM-PQVSVMGLGRMFWRREE 59  
DB 1 MTLRLLEDWCRGMDVNSQRTLLVWGPVNCDEAIEBETLQAAM-PQVSVMGLGRMFWRREE 60  
QY 60 NAKAALLELTGAVDYAAIPREMPKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDV 119  
DB 61 NRKVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSRLNEFLAGEGWTGEL 120  
QY 120 ARVLGFQNPPT--PTPG--PEMPAEMLNLYILDVNIQPLVESIWKRLTLFSGKHPRAWRG 175  
DB 121 SRAIGHENGSLDPEQGMIPEMWAPMLAQALE-ALQPALQCLKYKLVFSGRSPFEGEE 179  
QY 176 NFDPMLEHTNEVLEEQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 235  
DB 180 EFGRWMFHTTQMIKAMQVDPVDEKRRRLLESRLRGPAALDVIRLVKINNPLITVDCEQLALEE 239  
QY 236 VFGSVSSRDQAQKFLNTYQNPCKLSAVVIRLEPLLOKVVEKGAIDKONVQARLEQVI 295  
DB 240 VFGVTDNPRELVQKYLTTTQKDEKLSAVVIRLEPLLOKVVEKGAIDKONVQARLEQVI 299  
QY 296 AGANHSGAIRRQLWLTGAGEGPG 318  
DB 300 AGAVHK-TIRRELNPEDGPAFG 321

RESULT 6  
AAB43023  
ID AAB43023 standard; protein; 312 AA.  
XX





DR WPI; 2000-468119/41.  
 XX N-PSDB; AAA60837.  
 PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic  
 PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test  
 PT sample for the presence or absence of antibodies to a Ma family  
 PT polypeptide.  
 XX  
 PS Claim 48; Fig 9-10; 27pp; Japanese.  
 XX  
 CC The present invention describes a method for diagnosing a paraneoplastic  
 CC syndrome or neoplasm. The method comprises assessing a test sample for  
 CC the presence or absence of antibodies to a Ma family polypeptide (I). The  
 CC method is used to diagnose a paraneoplastic syndrome especially  
 CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or  
 CC neoplasm in an individual. The method diagnoses the neoplasm by assessing  
 CC antibodies to (I) preferably Ma1, which is indicative presence of breast  
 CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular  
 CC cancer and germ-cell tumours or Ma2, which is indicative of testicular  
 CC cancer, germ-cell tumour, and lung cancer. The present sequence is the  
 CC Ma5 protein as given in the present invention  
 XX  
 SQ Sequence 463 AA;  
 Query Match 44.3%; Score 766.5; DB 3; Length 463;  
 Best Local Similarity 50.2%; Pred. No. 3e-69;  
 Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;  
 QY 1 MAMTLLDWCRCMDVNSORTLLVWGIPVNCDEABIEETLQAAAPQVS-YRMLGRMFWRREE 59  
 DB 1 MPTLLQDWCRGEHLNTRCMLILGIPEDCGEDEFETLQACRHLGRYRVIGRMFRREE 60  
 QY 60 NAKAALLELTGADVAAIPREMPGKGVKVLFKPTSDAEFLRLHLFLAREGTVQDV 119  
 DB 61 NQAALLLELAQIDYALLPREIPGKGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDM 120  
 QY 120 ARVLGFQNPPTPGPMPAEMLVN--ILDNVIQPLVESIWKRLTLFSGKGHPRAWGNF 177  
 DB 121 NRVLGSDTNCAPRTVISPEFWTAAQLGAAVQPLLEQMLRELTVFSGNTISIPGALAF 180  
 QY 178 DPWLHTNEVLQVSDVEKRRRLMESLRGPAADVIRILKSNNPATITAECLKALEQVF 237  
 DB 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPALQVWSGLRASNASITVEECLAALQOVF 240  
 QY 238 GSVSSRDAQIKELNTYQNPCKLSAVYRLEPLQKVGKAIKDNVQNAQLAEQVAG 297  
 DB 241 GPVESHKIAQVLCXAYQGEAGEKVSFVIRLEPLQRAVENNVSRNVNQTRLKRVLSG 300  
 QY 298 ANHSGAIRQLWL 310  
 DB 301 ATLPLKLRDKLKL 313  
 RESULT 10  
 AAB42315  
 ID AAB42315 standard; protein; 463 AA.  
 XX  
 AC AAB42315;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF2079 polypeptide sequence SEQ ID NO:4159.  
 XX  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antiparotid; antiparkinsonian; nontropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antithyroid; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;

KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX Homo sapiens.  
 OS  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 XX 31-MAR-2000; 2000WO-US008621.  
 XX  
 PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinketsu RA, Leach M;  
 XX  
 DR WPI; 2000-602362/57.  
 DR N-PSDB; AAC76524.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 11; Page 3345-3347; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 CC antiparotid; antiparkinsonian; nontropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antithyroid; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotrophic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 SQ Sequence 463 AA;  
 Query Match 44.3%; Score 766.5; DB 3; Length 463;  
 Best Local Similarity 50.2%; Pred. No. 3e-69;  
 Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;  
 QY 1 MAMTLLDWCRCMDVNSORTLLVWGIPVNCDEABIEETLQAAAPQVS-YRMLGRMFWRREE 59  
 DB 1 MPTLLQDWCRGEHLNTRCMLILGIPEDCGEDEFETLQACRHLGRYRVIGRMFRREE 60  
 QY 60 NAKAALLELTGADVAAIPREMPGKGVKVLFKPTSDAEFLRLHLFLAREGTVQDV 119  
 DB 61 NQAALLLELAQIDYALLPREIPGKGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDM 120  
 QY 120 ARVLGFQNPPTPGPMPAEMLVN--ILDNVIQPLVESIWKRLTLFSGKGHPRAWGNF 177  
 DB 121 NRVLGSDTNCAPRTVISPEFWTAAQLGAAVQPLLEQMLRELTVFSGNTISIPGALAF 180  
 QY 178 DPWLHTNEVLQVSDVEKRRRLMESLRGPAADVIRILKSNNPATITAECLKALEQVF 237  
 DB 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPALQVWSGLRASNASITVEECLAALQOVF 240

QY 238 GSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAG 297  
 DB 241 GPVESHKIAQVLCXKAYQAGEKVSFVLRLEPLQRAVENNVSRNVNQTLLKRVLSG 300  
 QY 298 ANHSGAIRRQLWL 310  
 DB 301 ATLDPKLRDKLKL 313  
 RESULT 11  
 ID AAO16179  
 XX AAO16179 standard; protein; 452 AA.  
 AC AAO16179;  
 XX  
 DT 28-MAR-2003 (first entry)  
 XX  
 DE Human protein #5.  
 XX  
 KW Human; vaccine; adult whole brain; foetal whole brain; tonsil;  
 KW adult hippocampus; disease-associated SNP analysis; knockout mouse;  
 KW disease model mouse; cancer; neurological disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200299103-A1.  
 XX  
 PD 12-DEC-2002.  
 XX  
 PF 27-MAY-2002; 2002WO-JP005134.  
 XX  
 PR 04-JUN-2001; 2001JP-00168370.  
 PR 16-AUG-2001; 2001JP-00246915.  
 XX  
 XX (KAZU-) KAZUSA DNA RES INST FOUND.  
 PA (PROT-) PROTEIN EXPRESS CO LTD.  
 XX  
 XX Ohara O, Nagase T, Nakajima D;  
 XX  
 XX MPI; 2003-140622/13.  
 DR N-PSDB; AAL51207.  
 XX  
 PT DNA preferentially expressed in human adult and fetal brain tissue useful  
 PT for diagnosis, treatment and analysis of cancer and mental disorders.  
 XX  
 PS Claim 1; Page 56-60; 73pp; Japanese.  
 XX  
 CC The invention comprises the amino acid and coding sequences of seven  
 CC human proteins that are preferentially expressed in adult whole brain,  
 CC foetal whole brain, tonsil and adult hippocampus tissue. The DNA  
 CC sequences are useful for the analysis of disease-associated single  
 CC nucleotide polymorphisms and the production of knockout and human disease  
 CC model mice. The DNA and protein sequences of the invention are useful for  
 CC the prevention (vaccine) and treatment of cancer and neurological  
 CC disorders. The present amino acid sequence represents a human protein of  
 CC the invention  
 XX  
 SQ Sequence 452 AA;  
 Query Match 43.0%; Score 744; DB 6; Length 452;  
 Best Local Similarity 46.0%; Pred. No. 5.8e-67;  
 Matches 143; Conservative 64; Mismatches 100; Indels 4; Gaps 2;  
 QY 1 MAMTLLEDWCRGMDVNSQRTLLVWGIPVNCDEAEIETLQAM-PQVSYRMGLGMFWREE 59  
 DB 5 MALTLLDWCCKGMDMPKALLIVGIPMECSVEIQTIVKAGLQPLCAYRVLGRMFRRED 64  
 QY 60 NAKAALDELTCADVDAIIPREMPGKGGVWKVLFKPTSDAEFLERLHLFLAREGWTVQDV 119  
 DB 65 NAKAVFTELADVTNYTILPSHIPKGGGSEWVVKPRNPDDFELSRNLNFKDEGRSMTDV 124  
 QY 120 ARVLGFQNPFTTGPPEMPAEMNLNLDNVIQPLVESIWKRLTLFSGKGHPRAWRGNDP 179

DB 125 ARALGC---CSLPASLDAEVMPQVRSPLEPPKESMYRKLKVFSGTASPSGGETTFD 181  
 QY 180 WLEHTNEVLEEQVSDVEKRRLESLGPAADVIRILKSNPAITTAELCKALEQVFGS 239  
 DB 182 WLEQVTEIMPITWQSEVEKRRLESLGPAALSIMRVLOANDSITVECLDALKQIFGD 241  
 QY 240 VESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAG 299  
 DB 242 KEDFRASQRFLOTSPKIGEKVSTFLLRLEPLQKAVHKSPLSVRSTDMIRLKLILARVA 301  
 QY 300 HSGAIRRQLWL 310  
 DB 302 MTPALRGKLEL 312  
 RESULT 12  
 ID ADA54410  
 XX ADA54410 standard; protein; 399 AA.  
 AC ADA54410;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Human protein, SEQ ID 1978.  
 XX  
 KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
 KW inflammatory disease; osteoporosis; neurological disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1293569-A2.  
 XX  
 PD 19-MAR-2003.  
 XX  
 PF 21-MAR-2002; 2002EP-00006586.  
 XX  
 PR 14-SEP-2001; 2001JP-00328381.  
 PR 24-JAN-2002; 2002US-0350435P.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX  
 DR MPI; 2003-395539/38.  
 DR N-PSDB; ADA52771.  
 XX  
 CC New polynucleotides encoding full-length polypeptides, e.g. secretory  
 CC and/or membrane proteins, useful for developing medicines for diseases in  
 CC PT which the gene is involved, or as target molecules for gene therapy.  
 XX  
 PS Claim 14; SEQ ID NO 1978; 205pp; English.  
 XX  
 CC The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 XX  
 SQ Sequence 399 AA;  
 Query Match 35.9%; Score 620; DB 6; Length 399;  
 Best Local Similarity 40.9%; Pred. No. 2.5e-54;  
 Matches 137; Conservative 64; Mismatches 106; Indels 28; Gaps 6;  
 QY 1 MAMTLLEDWCRGMDVNSQRTLLVWGIPVNCDEAEIETLQAM-PQVSYRMGLGMFWREE 59  
 DB 1 MAYTMLQDWCRCWGVNARRLLILGIPEDCDDAEFQBSLEAALRPMGHFTVLGKAFREED 60



XX WPI; 2001-308781/32.

XX New isolated nucleic acid molecule encoding a human secreted protein is

XX used in preventing, treating or ameliorating a medical condition.

XX

XX Disclosure; Page 46; 519pp; English.

XX

XX AAD05121-AAD05203 represent cDNAs corresponding to 24 human secreted

XX protein genes, and AAE01232-AAE01311 represent the proteins they encode.

XX AAE01312-AAE01340 represent human secreted protein variants or fragments.

XX The secreted proteins and their genes are useful for preventing, treating

XX or ameliorating medical conditions, e.g., by protein or gene therapy.

XX Pathological conditions can be diagnosed by determining the amount of the

XX new protein in a sample or by determining the presence of mutations in

XX the new genes. Specific uses are described for each of the 24 genes,

XX based on the tissues in which they are most highly expressed, and include

XX developing products for the diagnosis or treatment of proliferative

XX disorders, cancer, tumours, foetal and developmental abnormalities,

XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune

XX diseases (e.g., rheumatoid arthritis), inflammation, allergies, disease),

XX neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),

XX cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,

XX psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,

XX angiogenic disorders, kidney disorders, gastrointestinal disorders,

XX pregnancy-related disorders, endocrine disorders, and infectious cell

XX proteins can also be used to aid wound healing and epithelial cell

XX proliferation, to prevent skin aging due to sunburn, to maintain organs

XX before transplantation, for supporting cell culture of primary tissues,

XX to regenerate tissues, to identify their cognate ligands or binding

XX partners, and in chemotaxis, and can be used as a food additive or

XX preservative to modify storage properties. Antibodies specific for a

XX protein of the invention can be used in alleviating symptoms associated

XX with the disorders mentioned above, and in diagnostic immunoassays e.g.,

XX radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The

XX present sequence represents a human secreted protein fragment referred to

XX in the disclosure of the invention

XX

XX SQ Sequence 280 AA;

Query Match 33.0%; Score 570.5; DB 4; Length 280;

Best Local Similarity 54.9%; Pred. No. 1.7e-49;

Matches 123; Conservative 35; Mismatches 57; Indels 9; Gaps 5;

QY 99 AEFLEHLFLAREGWTQDVAVLGFQNP--PTPG--PEMAEMLYLDNVIOPLVE 154

DB 32 AEFGR--FLAGEGTVGLSLRGHENGSLDPEQGMPEMWPAPMAQALE-ALQPALQ 87

QY 155 SIWYKRLTLFGKGHPRAWKGNFDPWLEHTNEVLEWQSDVDEKRRRLMESLRGPAADVI 214

DB 88 CLKYKLRVFGRESPEPEEGRWFWHTTQMIKAWQVDVDEKRRRLLESRLGPALDVI 147

QY 215 RIKSNNPALTTRACLEKALQEQVGSVSSRDAQIKFNTYQNGEKLAVVIEPLQK 274

DB 148 RVLIKNPLITVDCLQALEEVEGVTDNPRELQVKYLTYYQDEEKLAVVLRLEPLQK 207

QY 275 VVEKGALDKNNVQARLEQVIAGANHSGARIRROLWLITGAGEGPG 318

DB 208 LVORGAIERDANNQARDQVIAGAVHK-TIRRELNPEDGPAFG 250

RESULT 15

AAE12528

ID AAE12528 standard; protein; 283 AA.

XX

XX AAE12528;

XX

XX 02-NOV-2000 (first entry)

XX

XX Human Ma4 protein SEQ ID NO:11.

XX

XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;

XX paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;

XX

KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;

XX germ-cell tumour.

XX

XX Homo sapiens.

XX

XX JF2000146982-A.

XX

XX 26-MAY-2000.

XX

XX 10-NOV-1999; 99JP-00320171.

XX

XX 10-NOV-1999; 98US-00189527.

XX

XX (SLOK ) SLOAN KETTERING INST CANCER RES.

XX

XX WPI; 2000-468119/41.

XX

XX N-PSDB; AAE0836.

XX

XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic

XX encephalitis or neoplasm e.g. colon cancer comprising assessing a test

XX sample for the presence or absence of antibodies to a Ma family

XX polypeptide.

XX

XX Claim 48; Fig 7-8; 27pp; Japanese.

XX

XX The present invention describes a method for diagnosing a paraneoplastic

XX syndrome or neoplasm. The method comprises assessing a test sample for

XX the presence or absence of antibodies to a Ma family polypeptide (1). The

XX method is used to diagnose a paraneoplastic syndrome especially

XX paraneoplastic limbic encephalitis and/or brainstem encephalitis or

XX neoplasm in an individual. The method diagnoses the neoplasm by assessing

XX antibodies to (1) preferably Ma1, which is indicative presence of breast

XX cancer, colon cancer, parotid gland cancer, lung cancer, testicular

XX cancer and germ-cell tumours or Ma2, which is indicative of testicular

XX cancer, germ-cell tumour, and lung cancer. The present sequence is the

XX Ma4 protein as given in the present invention

XX

XX SQ Sequence 283 AA;

Query Match 32.7%; Score 565; DB 3; Length 283;

Best Local Similarity 47.0%; Pred. No. 6.3e-49;

Matches 117; Conservative 47; Mismatches 73; Indels 12; Gaps 4;

QY 83 GKGGVWKVLFKPPSTDAEFLEHLFLAREGWTQDVAVLGFQNPPTTP--GPEMPA 138

DB 3 GKGGVWKVLFKPPSTDAEFLEHLFLAREGWTQDVAVLGFQNPPTTP--GPEMPA 62

QY 139 EMLNYILDNVIOPLVESIWKRLTLFSGKGHPRAWKGNFDPWLEHTNEVLEWQSDVDEK 198

DB 63 HLIGQAWAHAPQPLL-PMYRKLURVFGSAVPAPEDSFEVWLEQATEIVKEWPTAEAK 121

QY 199 RRRLMESLRGPAADVIRILKSNNPALTTRACLEKALQEQVGSVSSRDAQIKFNTYQNGP 258

DB 122 KRWLAESLRGPAADVIRILKSNNPALTTRACLEKALQEQVGSVSSRDAQIKFNTYQNGP 181

QY 259 EKLAVVIEPLQKVKVEKALDKNNVQARLEQVIAGANHSGARIRROLWLITGAGEGPG 315

DB 182 EKVSAYVLRLETLRKAVKRAIPRRITADQVRLEQVMAGA---TLNQMLWCRLEKQK 237

QY 316 GPGPKPLSV 324

DB 238 GPPPSFLEL 246

Search completed; September 21, 2004, 13:35:12

Job time : 63.7447 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 76.2367 Seconds  
(without alignments)  
1361.621 Million cell updates/sec

Title: US-10-037-860-4  
Perfect score: 1729  
Sequence: 1 MAMTLEDWCRGMDVNSORT.....LTGAGEGPGPKPLSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25.\*  
1: sp archaea.\*  
2: sp bacteria.\*  
3: sp fungi.\*  
4: sp human.\*  
5: sp invertebrate.\*  
6: sp mammal.\*  
7: sp mhc.\*  
8: sp organelle.\*  
9: sp phase.\*  
10: sp plant.\*  
11: sp rodent.\*  
12: sp virus.\*  
13: sp vertebrate.\*  
14: sp unclassified.\*  
15: sp virus.\*  
16: sp bacteriap.\*  
17: sp archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	92.7	353	4 Q95144	Q95144 homo sapien
2	1602	92.7	353	4 Q8NG07	Q8NG07 homo sapien
3	1525	88.2	353	11 Q8VHZ4	Q8VHZ4 rattus norv
4	1507	87.2	353	11 Q9CYP2	Q9CYP2 mus musculu
5	1501	86.8	353	11 Q8C1C8	Q8C1C8 mus musculu
6	887.5	51.3	351	4 Q96BY2	Q96BY2 homo sapien
7	887.5	51.0	351	4 Q9HAS1	Q9HAS1 homo sapien
8	875.5	50.6	351	6 Q95K14	Q95K14 macaca fasc
9	870.5	50.3	352	11 Q9ERH6	Q9ERH6 mus musculu
10	772.5	45.8	466	11 Q8JZW8	Q8JZW8 mus musculu
11	770.5	44.9	364	6 Q9GMU3	Q9GMU3 macaca fasc
12	770.5	44.6	364	4 Q94959	Q94959 homo sapien
13	766.5	44.3	455	4 Q9H0A4	Q9H0A4 homo sapien
14	766.5	44.3	463	4 Q9UL41	Q9UL41 homo sapien
15	750	43.4	192	4 Q8ND90	Q8ND90 homo sapien
16	744	43.0	448	4 Q8NET3	Q8NET3 homo sapien

17	744	43.0	452	4 Q96PV4	Q96PV4 homo sapien
18	704	40.7	365	11 Q8BHK0	Q8BHK0 mus musculu
19	620	35.9	399	4 Q96A40	Q96A40 homo sapien
20	564	32.6	283	4 Q9UL42	Q9UL42 homo sapien
21	540	31.2	237	4 Q9H833	Q9H833 homo sapien
22	431	24.9	194	4 Q95145	Q95145 homo sapien
23	386.5	22.4	378	4 Q8N3H4	Q8N3H4 homo sapien
24	386.5	22.4	435	4 Q86V59	Q86V59 homo sapien
25	369.5	21.4	246	11 Q8C533	Q8C533 mus musculu
26	369.5	21.4	430	11 Q80VM8	Q80VM8 mus musculu
27	335.5	19.4	403	4 Q8TE36	Q8TE36 homo sapien
28	332.5	19.2	402	11 Q9CZA5	Q9CZA5 mus musculu
29	332	19.2	402	4 Q8N1C1	Q8N1C1 homo sapien
30	315.5	18.2	393	11 Q9DBI7	Q9DBI7 mus musculu
31	315.5	18.2	393	11 Q8VD24	Q8VD24 mus musculu
32	256.5	14.8	149	4 Q9UL43	Q9UL43 homo sapien
33	194.5	11.2	327	11 Q8VC32	Q8VC32 mus musculu
34	192	11.1	538	4 Q9H0W5	Q9H0W5 homo sapien
35	192	11.1	538	4 Q8TB26	Q8TB26 homo sapien
36	114.5	6.6	2037	5 Q22511	Q22511 caenorhabdi
37	111	6.4	555	16 Q82TR8	Q82TR8 nitrosomona
38	111	6.4	558	16 Q9JSW5	Q9JSW5 neisseria m
39	110.5	6.4	559	2 Q9F1V6	Q9F1V6 myxococcus
40	109.5	6.3	553	16 Q8PCN6	Q8PCN6 xanthomonas
41	109	6.3	652	16 Q7V5L3	Q7V5L3 prochloroco
42	108	6.2	558	16 Q9K1I2	Q9K1I2 neisseria m
43	107	6.2	555	16 Q8XVB8	Q8XVB8 raistonia s
44	106.5	6.2	503	16 Q8NPK8	Q8NPK8 corynebacte
45	106	6.1	518	16 Q8G6F5	Q8G6F5 bifidobacte

## ALIGNMENTS

### RESULT 1

Q95144 ID Q95144 PRELIMINARY; PRT; 353 AA.  
AC Q95144;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)  
DE 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Paraneoplastic neuronal antigen MAL.  
GN MAL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Cerebellum;  
RC MEDLINE=99158179; PubMed=10050892;  
RA Dalmay J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,  
RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frennier J.,  
RA Posner J.B., Rosenfeld M.R.;  
RT "Mal, a novel neuron- and testis-specific protein, is recognized by  
RT the serum of patients with paraneoplastic neurological disorders.";  
RL Brain 122:27-39(1999).  
[2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Cerebellum;  
RC Dalmay J., Rosenfeld M.R., Voltz R., Hoard R.;  
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF037364; AAD13810.3; -;  
DR Genew; HGNC:9158; FNMA1.  
DR GO; GO:0005737; C:cytoplasm; TAS.  
DR GO; GO:0005730; C:nucleolus; TAS.  
DR GO; GO:0007417; P:central nervous system development; TAS.  
DR GO; GO:0007283; P:spermatogenesis; TAS.  
SQ SEQUENCE 353 AA; 39800 MW; 3BB41691AE89AD3D CRC64;

Query Match 92.7%; Score 1602; DB 4; Length 353;  
Best Local Similarity 96.6%; Pred. No. 1.6e-120;  
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMGLGRMFWEEN 60  
 DB 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMGLGRMFWEEN 60  
 QY 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
 DB 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
 QY 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKHPRAWGNFDPW 180  
 DB 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKHPRAWGNFDPW 180  
 QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240  
 DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240  
 QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
 DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
 QY 301 SGAIRQLWLTGAGEGPP 319  
 DB 301 SGAIRQLWLTGAGEGPP 319

## RESULT 2

Q8NG07 ID Q8NG07 PRELIMINARY; PRT; 353 AA.  
 AC Q8NG07  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Paraneoplastic antigen.  
 GN PNMA1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schueller M.M., Jenne D.E., Schutze dit Belkner N., Hohlfeild R.,  
 RA Voltz R.;  
 RT "Mal (PNMA1).";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Prostate;  
 RA Strausberg R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF320308; AAN05100.1; -.  
 DR EMBL; BC039577; AAH39577.1; -.  
 SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;

Query Match 92.7%; Score 1602; DB 4; Length 353;  
 Best Local Similarity 96.6%; Pred. No. 1.6e-120;  
 Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMGLGRMFWEEN 60  
 DB 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMGLGRMFWEEN 60  
 QY 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
 DB 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
 QY 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKHPRAWGNFDPW 180  
 DB 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKHPRAWGNFDPW 180  
 QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240  
 DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240

QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
 DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
 QY 301 SGAIRQLWLTGAGEGPP 319  
 DB 301 SGAIRQLWLTGAGEGPP 319

## RESULT 3

Q8VHZ4 ID Q8VHZ4 PRELIMINARY; PRT; 353 AA.  
 AC Q8VHZ4;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Paraneoplastic onconeural protein MA1.  
 GN MA1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pellkofer H.L., Schubart A., Pagany M., Voltz R.D., Linington C.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF335505; AAL73196.1; -.  
 SQ SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEADA3 CRC64;

Query Match 88.2%; Score 1525; DB 11; Length 353;  
 Best Local Similarity 90.0%; Pred. No. 2.5e-114;  
 Matches 287; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMGLGRMFWEEN 60  
 DB 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVSYRMGLGRMFWEEN 60  
 QY 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
 DB 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120  
 QY 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKHPRAWGNFDPW 180  
 DB 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKHPRAWGNFDPW 180  
 QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240  
 DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240  
 QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
 DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
 QY 301 SGAIRQLWLTGAGEGPP 319  
 DB 301 SGAIRQLWLTGAGEGPP 319

## RESULT 4

Q9CYP2 ID Q9CYP2 PRELIMINARY; PRT; 353 AA.  
 AC Q9CYP2;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE 5730402C15RIK protein.  
 GN PNMA1 OR 5730402C15RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]



Db 121 SRALGHENSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKKLRVFSGRESPEPGEE 179

Qy 176 MPDPLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 235

Db 180 EFGRMFHTTQMIKAWQVDPVDEKRRRLLESRGPAADVIRILKSNPAITTAECLEKALEE 239

Qy 236 VFGSVSSSDAQIKFNTYQNPGEKLSAVYIRLEPLLQKVVEKGAIDKDNVQARLEQVI 295

Db 240 VFGVTDNPRELQVKYLTTHKDEKLSAVYIRLEPLLQKVVEKGAIDKDNVQARLDQVI 299

Qy 296 AGANHSAGAIRRQLWLTGAGEGPG 318

Db 300 AGAVHK-TIRRELNPEDGPAPG 321

RESULT 7

Q9HAS1 PRELIMINARY; PRT; 351 AA.

AC Q9HAS1;

DT 01-MAR-2001 (TREMELrel. 16, Created)

DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)

DE MAP-1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Cerebellum;

RX MEDLINE=21264738; PubMed=11060313;

RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C., Yu V.C.;

RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that associates with Bax through its Bcl-2 homology domains.";

RL J. Biol. Chem. 276:2802-2807(2001).

DR EMBL; AF305550; AAG31786.1; -.

SQ SEQUENCE 351 AA; 39521 MW; 8736401AD8A24EAD CRC64;

Query Match 51.0%; Score 882.5; DB 4; Length 351;

Best Local Similarity 55.1%; Pred. No. 1.2e-62;

Matches 178; Conservative 53; Mismatches 85; Indels 7; Gaps 5;

Qy 1 MAMTLLDWCRCMDVNSQRTLLVWGPVNCDEAEIETLQAA-POVSYRMLGRMFWRDE 59

Db 1 MTLRLLEDWCRCMDMNPRAKLLIAGISQSCSVAEIEALQAGLAPLGEYRLIGRMFRRDE 60

Qy 60 NAKAALLELTGADVAAIAPREMPGGVGVKVPKPTSDAEFLERHLFLAREGTVQDV 119

Db 61 NRKVALVGLTAETSHALVPEIPIKGGIWRVIFKPPDPDNTFLSRNEFLAGEGTVGEL 120

Qy 120 ARVLGFQNPPTPG--PMPAEMLYILDNVIOPLVESIWYKRLTLFSGKGHP 175

Db 121 SRALGHENSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKKLRVFSGRESPE 179

Qy 176 MPDPLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 235

Db 180 EFGRMFHTTQMIKAWQVDPVDEKRRRLLESRGPAADVIRILKSNPAITTAECLEKALEE 239

Qy 236 VFGSVSSSDAQIKFNTYQNPGEKLSAVYIRLEPLLQKVVEKGAIDKDNVQARLEQVI 295

Db 240 VFGVTDNPRELQVKYLTTHKDEKLSAVYIRLEPLLQKVVEKGAIDKDNVQARLDQVI 299

Qy 296 AGANHSAGAIRRQLWLTGAGEGPG 318

Db 300 AGAVHK-TIRRELNPEDGPAPG 321

RESULT 8

Q9SK14 PRELIMINARY; PRT; 351 AA.

ID Q9SK14

AC Q9SK14;

DT 01-DEC-2001 (TREMELrel. 19, Created)

DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Hypothetical protein.

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

OC Cercopitheidae; Macaca.

OX NCBI\_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Temporal Cortex;

RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K.;

RT "Isolation of full-length cDNA clones from macaque brain cDNA libraries.";

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB060854; BAB46873.1; -.

KW Hypothetical protein.

SQ SEQUENCE 351 AA; 39623 MW; C753054496A6FFB3 CRC64;

Query Match 50.6%; Score 875.5; DB 6; Length 351;

Best Local Similarity 54.3%; Pred. No. 4.2e-62;

Matches 178; Conservative 51; Mismatches 82; Indels 17; Gaps 5;

Qy 1 MAMTLLDWCRCMDVNSQRTLLVWGPVNCDEAEIETLQAA-POVSYRMLGRMFWRDE 59

Db 1 MTLRLLEDWCRCMDMNPRAKLLIAGISQSCSVAEIEALQAGLAPLGEYRLIGRMFRRDE 60

Qy 60 NAKAALLELTGADVAAIAPREMPGGVGVKVPKPTSDAEFLERHLFLAREGTVQDV 119

Db 61 NRKVALVGLTAETSHALVPEIPIKGGIWRVIFKPPDPDNTFLSRNEFLAGEGTVGEL 120

Qy 120 ARVLGFQNPPTPG--PMPAEMLYILDNVIOPLVESIWYKRLTLFSGKGHP 170

Db 121 TRALAHEN-----GSLDLEQGMIPENWAPMLAQALE-ALQPALQCLKYKKLRVFSGRE 174

Qy 171 RAWGNFDPWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLE 230

Db 175 EPGEEFGRMFHTTQMIKAWQVDPVDEKRRRLLESRGPAADVIRILKSNPAITTAECLE 234

Qy 231 KALEQVFGSVSSSDAQIKFNTYQNPGEKLSAVYIRLEPLLQKVVEKGAIDKDNVQAR 290

Db 235 QALEEVFGVTDNPRELQVKYLTTHKDEKLSAVYIRLEPLLQKVVEKGAIDKDNVQAR 294

Qy 291 LEQVIAGANHSAGAIRRQLWLTGAGEGPG 318

Db 295 LDQVIAGAVHK-TIRRELNPEDGPAPG 321

RESULT 9

Q9ERH6 PRELIMINARY; PRT; 352 AA.

ID Q9ERH6

AC Q9ERH6;

DT 01-MAR-2001 (TREMELrel. 16, Created)

DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)

DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)

DE MAP-1 (4930435G24RIK protein) (RIKEN CDNA 4930435G24 gene).

GN MOAP1 OR 4930435G24RIK.

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C., Yu V.C.;

RT "MAP-1: a novel pro-apoptotic protein containing a BH3-like motif that associates with Bax through its Bcl-2 homology domains.";

RL J. Biol. Chem. 0:0-0(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;



Query Match 44.9%; Score 776.5; DB 6; Length 364;  
 Best Local Similarity 48.2%; Pred. No. 48-54; 95; Indels 13; Gaps 5;  
 Matches 160; Conservative 64; Mismatches 95; Indels 13; Gaps 5;

1 MAMTLEDCRCMDVNSQRTLLVWGIPIVNCDAEIEETLQAAMPQV-SYRMLGRMFWREE 59  
 1 MALALLEDCRIMSDEQKSLMVTGIPADFEAEIOEVLQETLKSIGSYRLLGKIFRQOE 60

60 NAKAALLELTGAVDYAAIPREMPGKGWVKVLPKPTSDAEFLERLHLFLAREGWTVDV 119  
 61 NANAVLLELLEDVSAIPSEVQGGWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGM 120

120 ARVLGFQNPPTP-----GPEMAEMLNYLDNVIOPLVESIWKRLTLFSGKGHPRAWRG 175  
 121 FRALGEGSPATVPCISPPELLAHLLGQMAHAPQLL-PMYRKLRVFGSAVPAPEE 179

176 NPDPLWHTNEVLEWQVSDVEKRRRLMESLGRPAADVIRILKSNPNPAITTAECLEKALQ 235  
 180 PFEVLEQATEIVKEWVPVTEAEKKRWLAESLRGPAALDLMHIVQADNPSISVEECLEAPKQ 239

236 VFGSVESRDAQIKFLNTYQNGEKLISAVIRLEPLLOKVVEKGALDKDNNVQARLEQVI 295  
 240 VFGSLESRTAQVRLKTYQEGEKVSAYVLRLETLRLRAVEKRAIPRIADQVRLQV 299

296 AGANSHGAIRROLWL---TGAGEGPGPKPLSV 324  
 300 AGA-----TLNQMLWCRLRELKDGPPPNFLEL 327

## RESULT 12

094959 PRELIMINARY; PRT; 364 AA.

AC Q94959  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein KIAA0883 (Paraneoplastic associated brain-testis-cancer antigen).  
 GN KIAA0883.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99156230; PubMed=10048485;  
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."  
 RT DNA Res. 5:355-364 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Richen J.G., Dalmau J., Wade D., Rosenfeld M.R.;  
 RT "Characterization of a Brain-Testis-Cancer Antigen."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB020690; BAA74906.1; -  
 DR EMBL; AF286487; AAG28165.1; -  
 DR Genew; HGNC:9159; PNMA2.  
 DR InterPro; IPR005162; Retrotrans\_gag.  
 DR Pfam; PF03732; Retrotrans\_gag; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0E93 CRC64;

Query Match 44.6%; Score 770.5; DB 4; Length 364;  
 Best Local Similarity 47.6%; Pred. No. 1.2e-53;  
 Matches 158; Conservative 65; Mismatches 96; Indels 13; Gaps 5;

1 MAMTLEDCRCMDVNSQRTLLVWGIPIVNCDAEIEETLQAAMPQV-SYRMLGRMFWREE 59  
 1 MALALLEDCRIMSDEQKSLMVTGIPADFEAEIOEVLQETLKSIGSYRLLGKIFRQOE 60

60 NAKAALLELTGAVDYAAIPREMPGKGWVKVLPKPTSDAEFLERLHLFLAREGWTVDV 119  
 61 NANAVLLELLEDVSAIPSEVQGGWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGM 120

120 ARVLGFQNPPTP-----GPEMAEMLNYLDNVIOPLVESIWKRLTLFSGKGHPRAWRG 175  
 121 FRALGEGSPATVPCISPPELLAHLLGQMAHAPQLL-PMYRKLRVFGSAVPAPEE 179

176 NPDPLWHTNEVLEWQVSDVEKRRRLMESLGRPAADVIRILKSNPNPAITTAECLEKALQ 235  
 180 PFEVLEQATEIVKEWVPVTEAEKKRWLAESLRGPAALDLMHIVQADNPSISVEECLEAPKQ 239

236 VFGSVESRDAQIKFLNTYQNGEKLISAVIRLEPLLOKVVEKGALDKDNNVQARLEQVI 295  
 240 VFGSLESRTAQVRLKTYQEGEKVSAYVLRLETLRLRAVEKRAIPRIADQVRLQV 299

296 AGANSHGAIRROLWL---TGAGEGPGPKPLSV 324  
 300 AGA-----TLNQMLWCRLRELKDGPPPNFLEL 327

Db 1 MALALLEDCRIMSDEQKSLMVTGIPADFEAEIOEVLQETLKSIGSYRLLGKIFRQOE 60

QY 60 NAKAALLELTGAVDYAAIPREMPGKGWVKVLPKPTSDAEFLERLHLFLAREGWTVDV 119

Db 61 NANAVLLELLEDVSAIPSEVQGGWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGM 120

QY 120 ARVLGFQNPPTP-----GPEMAEMLNYLDNVIOPLVESIWKRLTLFSGKGHPRAWRG 175

Db 121 FRALGEGSPATVPCISPPELLAHLLGQMAHAPQLL-PMYRKLRVFGSAVPAPEE 179

QY 176 NPDPLWHTNEVLEWQVSDVEKRRRLMESLGRPAADVIRILKSNPNPAITTAECLEKALQ 235

Db 180 PFEVLEQATEIVKEWVPVTEAEKKRWLAESLRGPAALDLMHIVQADNPSISVEECLEAPKQ 239

QY 236 VFGSVESRDAQIKFLNTYQNGEKLISAVIRLEPLLOKVVEKGALDKDNNVQARLEQVI 295

Db 240 VFGSLESRTAQVRLKTYQEGEKVSAYVLRLETLRLRAVEKRAIPRIADQVRLQV 299

QY 296 AGANSHGAIRROLWL---TGAGEGPGPKPLSV 324

Db 300 AGA-----TLNQMLWCRLRELKDGPPPNFLEL 327

## RESULT 13

Q9H0A4 PRELIMINARY; PRT; 455 AA.

AC Q9H0A4  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN DKFZP434K225.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 RA Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,  
 RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,  
 RA Wambutt R., Korn B., Klein M., Poustka A.;  
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and Analysis of 500 Novel Complete Protein Coding Human cDNAs."  
 RT Genome Res. 11:422-435 (2001).  
 RL EMBL; AL136878; CAB66812.1; -  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; zf-CCHC; 1.  
 DR PRINTS; PR00939; C2HCZNFINGER.  
 DR SMART; SM00343; Znf\_C2HC; 1.  
 DR PROSITE; P550158; ZF\_CCHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 455 AA; 51514 MW; 9AC8CC06685A8956 CRC64;

Query Match 44.3%; Score 766.5; DB 4; Length 455;  
 Best Local Similarity 50.2%; Pred. No. 3.5e-53;  
 Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;

1 MAMTLEDCRCMDVNSQRTLLVWGIPIVNCDAEIEETLQAAMPQV-SYRMLGRMFWREE 59  
 1 MPTLLQDCRCRGLHNTTRCMLTILGIPEDGGEFEETLQEAACRHLGRYVIGRMFEREE 60

60 NAKAALLELTGAVDYAAIPREMPGKGWVKVLPKPTSDAEFLERLHLFLAREGWTVDV 119  
 61 NAOAILLELAQDIDYALLPREIPFGKGPWEVIVKPNSDGFEFLNRLNRFLEEBRTVSDM 120

120 ARVLGFQNPPTP-----GPEMAEMLNYLDNVIOPLVESIWKRLTLFSGKGHPRAWRG 177  
 121 NRVLGSDTNCSPRVISPEFWTWAQTLGNAVQLLEQMLYRELRFVSGNTISIPGALAF 180

```

QY 178 DPWLEHTNEVLEEMQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQOVF 237
Db 181 DAWLEHTTEMLQWQVPEGEKRRLEMECLRGPAQVSGLRASNASITVEECALAQOVF 240

QY 238 GSVESRRDAQIKFLNTYONPGEKLSAYVIRLEPLLOKVEKGAIDKDNVNQARLEQVIAG 297
Db 241 GPVESHKIAQVKLCAYQAGEKVSFVLRLEPLLOQRAVENNVVSRNNVQTRLKRVLSG 300

QY 298 ANHSGAIRROLWL 310
Db 301 ATLPLDKLRDKLKL 313

RESULT 14
Q9UL41 PRELIMINARY; PRT; 463 AA.
AC Q9UL41;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Paraneoplastic neuronal antigen MA3.
GN MA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J.,
RA Posner J.B., Dalmay J.;
RT "Identification of a novel cancer testis brain antigen using serum
RT antibodies from patients with testicular tumors and paraneoplastic
RT limbic encephalomyelitis.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083116; AAF05627.1; -.
DR GO; HGNC:18742; PNMA3.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001878; Znf CCHC.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf C2HC; 1.
DR PROSITE; PS50158; ZF_CCHC; 1.
SQ SEQUENCE 463 AA; 52376 MW; 0843BF6601D8557C CRC64;

Query Match 44.3%; Score 766.5; DB 4; Length 463;
Best Local Similarity 50.2%; Pred. No. 3.6e-53;
Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;

QY 1 MAMTLLEDWCRGMDVNSQRTLLVWGIPVNCDEABIEETLQAAMPQVS-YRMLGRMFWRREE 59
Db 1 MPLTLQDWCRGHEHLNTRRCMLIIGIPEDCGEDFEETLQEAHRHGRYRVIGRMFRREE 60

QY 60 NAKAALLELGADVYAIPREMPQKGVKWKFKPPTSDAEFLRLHLFLAREGWTQDV 119
Db 61 NQAAILLELAQDIDYALLPREIPGKGPWEIVKPRNSDGEFLNRLNKFLEERRTVSDM 120

QY 120 ARVLGFQNPPTPGPMPAELNY--ILDNVIQPLVESIWKRLTLFSGKGHPRAWGNF 177
Db 121 NRVLGSNTNCSAPRVTSPEFWTQAQTLGAQVPLLEQMLYRELRLVFSGNTISIPGALAF 180

QY 178 DPWLEHTNEVLEEMQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQOVF 237
Db 181 DAWLEHTTEMLQWQVPEGEKRRLEMECLRGPAQVSGLRASNASITVEECALAQOVF 240

QY 238 GSVESRRDAQIKFLNTYONPGEKLSAYVIRLEPLLOKVEKGAIDKDNVNQARLEQVIAG 297
Db 241 GPVESHKIAQVKLCAYQAGEKVSFVLRLEPLLOQRAVENNVVSRNNVQTRLKRVLSG 300

QY 298 ANHSGAIRROLWL 310
Db 301 ATLPLDKLRDKLKL 313

```

```

RESULT 15
Q8ND90 PRELIMINARY; PRT; 192 AA.
AC Q8ND90;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP434N171.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Poustka A., Wellenreuther R., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834327; CAD38995.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 192 AA; 21456 MW; 5DB17CF6A760A628 CRC64;

Query Match 43.4%; Score 750; DB 4; Length 192;
Best Local Similarity 93.7%; Pred. No. 2.2e-52;
Matches 148; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 162 TLFSGKHPRAWGNFDPWLEHTNEVLEEMQVSDVEKRRRLMESLRGPAADVIRILKSN 221
Db 1 TLFSGRDIPGPGERTFDPWLEHTNEVLEEMQVSDVEKRRRLMESLRGPAADVIRILKSN 60

QY 222 PAITTAECLEKALQOVFGSVESRRDAQIKFLNTYONPGEKLSAYVIRLEPLLOKVEKAI 281
Db 61 PAITTAECLEKALQOVFGSVESRRDAQIKFLNTYONPGEKLSAYVIRLEPLLOKVEKAI 120

QY 282 DKDNVQARLEQVIAGANHSGAIRROLWLTGAGSGP 319
Db 121 DKDNVQARLEQVIAGANHSGAIRROLWLTGAGSGP 158

```

Search completed: September 21, 2004, 13:39:30  
Job time : 79.2367 secs

Blank

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 ; Search time 8.96902 Seconds  
(without alignment)  
1910.028 Million cell updates/sec

Title: US-10-037-860-4

Perfect score: 1729

Sequence: 1 MAMTLLEDWCRGMDVNSQRT.....LTGAGEGPGPKPLSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	92.7	353	1	PMAL HUMAN
2	1525	88.2	353	1	PMAL RAT
3	1501	86.8	353	1	PMAL MOUSE
4	887.5	51.3	351	1	MOP1 HUMAN
5	875.5	50.6	351	1	MOP1 MACFA
6	870.5	50.3	352	1	MOP1 MOUSE
7	776.5	44.9	364	1	PMAL2 MACFA
8	770.5	44.6	364	1	PMAL2 HUMAN
9	704	40.7	365	1	PMAL2 MOUSE
10	104.5	6.0	556	1	YJJK HAEIN
11	103	6.0	935	1	KINH SYNRA
12	101	5.8	488	1	SYFA AERPE
13	99.5	5.8	2564	1	SPCQ HUMAN
14	98.5	5.7	424	1	FLAI ARATH
15	98	5.7	513	1	YM26 MYCTU
16	97.5	5.6	759	1	PARC CAUCR
17	96	5.6	468	1	EFLA HYDAT
18	95.5	5.5	645	1	SYM CLOPE
19	92.5	5.3	464	1	KTR4 YEAST
20	92.5	5.3	2431	1	POLN SFV
21	91.5	5.3	378	1	B3GA HUMAN
22	90	5.2	404	1	CPXJ SACER
23	90	5.2	1660	1	VIT6 OSCBR
24	89	5.1	646	1	THIC VIBPA
25	89	5.1	1755	1	PEPL MOUSE
26	88.5	5.1	552	1	THSB SULTO
27	88.5	5.1	808	1	SYFB SYNFP
28	88	5.1	640	1	MYB BOVIN
29	87	5.0	461	1	EFLA APINE
30	87	5.0	600	1	BIR3 MOUSE
31	87	5.0	693	1	ERG_FUSNN
32	87	5.0	990	1	GOA2 HUMAN
33	87	5.0	6885	1	SNEZ2_HUMAN

34	86.5	5.0	605	1	VAA2 TREPA	O83541 treponema p
35	86.5	5.0	1155	1	IF2P METJA	Q57710 methanococ
36	86.5	5.0	1364	1	KINI_CHICK	Q90631 gallus gall
37	86	5.0	316	1	FLGJ_SALTY	P15931 salmoneilla
38	86	5.0	364	1	GUNA_RUMAL	P23660 ruminococu
39	86	5.0	424	1	PUR2 AQUAE	O66949 aquifex aeo
40	86	5.0	554	1	YJJK_ECOLI	P37797 escherichia
41	86	5.0	727	1	MUTB_PROFR	P11653 propionibac
42	85	4.9	926	1	RGP1_CAEEL	P34342 caenorhabdi
43	84.5	4.9	495	1	THSB_SULAC	Q9V2T4 sulfolobus
44	84.5	4.9	546	1	DX56_MOUSE	Q9D0R4 mus musculu
45	84.5	4.9	641	1	MYB_CHICK	P01103 gallus gall

RESULT 1

PMAL\_HUMAN STANDARD; PRT; 353 AA.

AC Q8ND50; Q85144; Q8NG07;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Paraneoplastic antigen Mal (Neuron- and testis-specific protein 1)  
 DE (37 kDa neuronal protein).  
 GN PMAL OR MAL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RX [1]  
 RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=99158179; PubMed=10050892;  
 RA Dalmaz J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,  
 RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Premier J.,  
 RA Posner J.B., Rosenfeld M.R.;  
 RT "Mal, a novel neuron- and testis-specific protein, is recognized by  
 RT the serum of patients with paraneoplastic neurological disorders.";  
 RL Brain 122:27-39(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Schueller M.M., Jenne D.E., Schutze dit Belkner N., Hohlfield R.,  
 RA Voltz R.;  
 RT "Mal (PNMAL).";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Prostate;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 162-353 FROM N.A.

RC TISSUE=Testis;  
RA Poustka A., Wellenreuther R., Mewes H.-W., Weill B., Wiemann S.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar. In tumor cells, it is  
CC cytoplasmic.  
CC -!- TISSUE SPECIFICITY: Testis and brain specific. In some patients  
CC suffering from cancers, it is also specifically expressed by the  
CC paraneoplastic tumor cells.  
CC -!- MISCELLANEOUS: Antibodies against PNMA1 are present in sera from  
CC patients suffering of paraneoplastic neurological disorders.  
CC -!- SIMILARITY: Belongs to the PNMA family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF037364; RAD13810.3; -;  
CC EMBL; AF320308; AAN05100.1; -;  
CC EMBL; BC039577; AAH39577.1; -;  
CC EMBL; AL834327; CAD38995.1; -;  
CC Genew; HGNC:9158; PNMA1.  
CC MIM: 604010; -;  
CC GO; GO:0005737; Cytoplasm; TAS.  
CC GO; GO:0005730; Cnucleolus; TAS.  
CC GO; GO:0007417; P:central nervous system development; TAS.  
CC GO; GO:0007283; P:spermatogenesis; TAS.  
CC Antigen; Tumor antigen; Nuclear protein.  
CC FT DOMAIN 335 341 POLY-GLU.  
CC SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;  
Query Match 92.7%; Score 1602; DB 1; Length 353;  
Best Local Similarity 96.6%; Pred. No. 5.6e-119;  
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 1 MAMTLLDRCGMDVNSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMLGRMFREEN 60  
Db 1 MAMTLLDRCGMDVNSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMLGRMFREEN 60  
QY 61 AKAALELTGAVDYAAIPREMPGKGVKWLFPKPTSDAEFLRLHLFLAREGWTVDVA 120  
Db 61 AKAALELTGAVDYAAIPREMPGKGVKWLFPKPTSDAEFLRLHLFLAREGWTVDVA 120  
QY 121 RVLFQNFPTPGPEMPAEMNLVLDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180  
Db 121 RVLFQNFPTPGPEMPAEMNLVLDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180  
QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPALTAAECLKALEQVFGSV 240  
Db 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPALTAAECLKALEQVFGSV 240  
QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPFLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
Db 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPFLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
QY 301 SGAIRQLMTGAGEGPP 319  
Db 301 SGAIRQLMTGAGEGPP 319

## RESULT 2

PNMA1\_RAT  
ID PNMA1\_RAT STANDARD; PRT; 353 AA.  
AC Q8VHZ4;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Paraneoplastic antigen Mal homolog.  
GN PNMA1 OR MAL.  
OS Rattus norvegicus (Rat).

OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RX MEDLINE=99158179; PubMed=10050892;  
RA Dalmat J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,  
RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frenn J.,  
RA Posner J.B., Rosenfeld M.R.;  
RT 'Mal, a novel neuron- and testis-specific protein, is recognized by  
RT the serum of patients with paraneoplastic neurological disorders.';  
RL Brain 122:27-39(1999).  
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).  
CC -!- TISSUE SPECIFICITY: Testis and brain specific.  
CC -!- SIMILARITY: Belongs to the PNMA family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF335505; AAL73196.1; -;  
CC Nuclear protein.  
CC FT DOMAIN 335 341 POLY-GLU.  
CC SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEDA3 CRC64;  
Query Match 88.2%; Score 1525; DB 1; Length 353;  
Best Local Similarity 90.0%; Pred. No. 6.6e-113;  
Matches 287; Conservative 15; Mismatches 17; Indels 0; Gaps 0;  
QY 1 MAMTLLDRCGMDVNSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMLGRMFREEN 60  
Db 1 MAMTLLDRCGMDVNSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMLGRMFREEN 60  
QY 61 AKAALELTGAVDYAAIPREMPGKGVKWLFPKPTSDAEFLRLHLFLAREGWTVDVA 120  
Db 61 AKAALELTGAVDYAAIPREMPGKGVKWLFPKPTSDAEFLRLHLFLAREGWTVDVA 120  
QY 121 RVLFQNFPTPGPEMPAEMNLVLDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180  
Db 121 RVLFQNFPTPGPEMPAEMNLVLDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180  
QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPALTAAECLKALEQVFGSV 240  
Db 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPALTAAECLKALEQVFGSV 240  
QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPFLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
Db 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPFLQKVVEKGAIDKDNVNQARLEQVIAGANH 300  
QY 301 SGAIRQLMTGAGEGPP 319  
Db 301 SGAIRQLMTGAGEGPP 319

## RESULT 3

PNMA1\_MOUSE  
ID PNMA1\_MOUSE STANDARD; PRT; 353 AA.  
AC Q8C1C8; Q9CVP2;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Paraneoplastic antigen Mal homolog.  
GN PNMA1 OR MAL.  
OS Mus musculus (Mouse).  
OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nixaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schrim L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Guscinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertele G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,  
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang J., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).  
 CC -!- SIMILARITY: Belongs to the PNA family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AK017476; BAB30762.1; -;  
 DR EMBL; AK028331; BAC25885.1; -;  
 DR MGD; MGI:2180564; Pnmal.  
 KW Nuclear protein.  
 FT DOMAIN 336 341 POLY-GLU.  
 FT CONFLICT 136 136 T -> M (IN REF. 1; BAB30762).  
 SQ SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;  
 Query Match 86.8%; Score 1501; DB 1; Length 353;  
 Best Local Similarity 88.7%; Pred. No. 5.1e-111;  
 Matches 283; Conservative 15; Mismatches 21; Indels 0; Gaps 0;  
 Qy 1 MAMTLEDCRCGMDVNSORTLLVWGIPVNCDEAEIEETLQAMPQVSYRLGRMFWRREN 60  
 Db 1 MAMTLEDCRCGMDVNSQALLVWGIPVNCDETEIEETLQAMPQVSYRLGRMFWRREN 60  
 Qy 61 AKAAELLELTGADVAAAPREMPGKGGVWVFKPPTSDAEFLERLHLFLAREGWTVDQVA 120  
 Db 61 AKAAELLELTGADVSLIPREMPGKGGVWVFKPPTSDAEFLERLHLFLAREGWTVDQVA 120  
 Qy 121 RVLGQNPPTPGMPAEMNLNLDNVIQPLVESIWKRLTLPSGKHPRAWRGNEPDW 180  
 Db 121 RVLGQNPAPAGPPTPAEMNLNLDNVIQPLVESIWKRLTLPSGKHPRAWRGNEPDW 180  
 Qy 181 LEHTNEVLEEQVSDVERRRRLMESLRGPAADVIRILKSNPNTATBCLKALEQVFGSV 240  
 Db 181 LEHSNEIIEEQVSDIEKRRRLMESLRGPAADVIRILKSNNTATBCLKALEQVFGSV 240  
 Qy 241 ESSRDAQIKFNTYQNPGEKLSAVYIRLEPLQKVEKGAIDKDNVNQARLEQVIAGNH 300  
 Db 241 ESSRDAQVRFNTYQNPGEKLSAVYIRLEPLQKVEKGAIDKDNVNQARLEQVIAGNH 300

Qy 301 SCATRRQLWLTGAGEGPGP 319  
 Db 301 SCALRRQLWLAGAERGAP 319  
 RESULT 4  
 MOPL HUMAN  
 ID MOPL HUMAN STANDARD; PRT; 351 AA.  
 AC Q96BY2; Q9H833; Q9HAS1;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DE Modulator of apoptosis 1 (MAP-1) (Paraneoplastic antigen MA4).  
 GN MOAP1 OR PMA4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF LEU-120;  
 RP 120-LEU-ARG-127 AND 125-GLY--GLU-127.  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=21264738; PubMed=11060313;  
 RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevoit M., Ang K.C.,  
 RA Yu V.C.;  
 RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that  
 RT associates with Bax through its Bcl-2 homology domains";  
 RL J. Biol. Chem. 276:2802-2807(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Retinoblastoma;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,  
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yanamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDD human cDNA sequencing project";  
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22386257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.B.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Mediates caspase-dependent apoptosis.  
 CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX.  
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with higher levels in  
 CC heart and brain.  
 CC -!- DOMAIN: The BH3-like domain is required for association with BAX  
 CC and for mediating apoptosis. The three BH domains (BH1, BH2, and  
 CC BH3) of BAX are all required for mediating protein-protein  
 CC interaction.

CC -!- SIMILARITY: Belongs to the PNMA family.  
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a  
 CC frameshift in position 102.

CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AF305550; RA031786.1; --  
 CC EMBL: AK024029; BAB14788.1; ALT\_SEQ.  
 CC EMBL: BC015044; AAH15044.1; --  
 CC Genew; HGNC:16658; MOAP1.  
 CC Apoptosis. 120 127 BH3-LIKE.  
 CC MUTAGEN 120 127 MISSING: ABROGATED INTERACTION WITH BAX,  
 CC RESULTING IN A NONAPOPTOTIC PROTEIN.  
 CC MUTAGEN 120 120 L->E: WEAKENED INTERACTION WITH BAX,  
 CC RESULTING IN A NONAPOPTOTIC PROTEIN.  
 CC MUTAGEN 125 127 GHE->VLA: ABROGATED INTERACTION WITH BAX,  
 CC RESULTING IN A NONAPOPTOTIC PROTEIN.  
 CC CONFLICT 244 244 T -> A (IN REF. 2).  
 CC CONFLICT 258 258 Y -> H (IN REF. 2).  
 CC CONFLICT 259 259 Q -> H (IN REF. 1).  
 CC SEQUENCE 351 AA; 39512 MW; 5310142AC02B563C CRC64;

Query Match 51.3%; Score 887.5; DB 1; Length 351;  
 Best Local Similarity 55.4%; Pred. No. 1.2e-62;  
 Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;

QY 1 MAMTLLDWCRCNDVNSQRTLLVWGIPIVNCDAEIEETLQAAM-PQVSYRMGLGRMFWREE 59  
 DB 1 MTRLLDWCRCNDVNSQRTLLVWGIPIVNCDAEIEETLQAAM-PQVSYRMGLGRMFWREE 60  
 QY 60 NAKAALLELTGAVDYAAIPREMPGKGVWVKVLPKPTSDAEFLERLHLFLAREGVTVDV 119  
 DB 61 NRKVALVGLTAETSHALVPEKIPGKGIWRVIFKPPDPNTFLSRLNEFLAGEGVTGEL 120  
 QY 120 ARVLGFQNPFT-PTPG--PEMPAEMLYLDNVIQPLVESIWKRLTLFSGKHPRAWRG 175  
 DB 121 SRALCHENGSLDPEQGMIPEMWAPMLAQALE-ALQALQCLYKLVFSGRSPGEE 179  
 QY 176 NFDPLWHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 235  
 DB 180 EFGRMWFTTQMIKAWQVDPVEKRRRLMESLRGPAADVIRILKSNPAITTVDECLQALEE 239  
 QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAIDKDNVNQARLEQVI 295  
 DB 240 VFGVTNPRELQVKYLTYYQKDEKLSAYVIRLEPLLOKVEKGAIDKDNVNQARLDQVI 299  
 QY 296 AGANHSAGAIRRQLWLTGAGEGPG 318  
 DB 300 AGAVHK-TIRRELNPEDGPAPG 321

## RESULT 5

MOPL\_MACFA STANDARD; PRT; 351 AA.  
 AC Q95KT4;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Modulator of apoptosis 1 (MAP-1).  
 GN MOAP1.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OC NCBI\_TaxID=9541;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Temporal Cortex;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Mediates caspase-dependent apoptosis (By similarity).  
 CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).  
 CC -!- DOMAIN: The BH3-like domain is required for association with BAX  
 CC and for mediating apoptosis. The three BH domains (BH1, BH2, and  
 CC BH3) of BAX are all required for mediating protein-protein  
 CC interaction (By similarity).  
 CC -!- SIMILARITY: Belongs to the PNMA family.

CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AB060854; BAB46873.1; --  
 CC Apoptosis. 120 127 BH3-LIKE.  
 CC SITE 336 339 POLY-GLU.  
 CC DOMAIN 351 AA; 39623 MW; C7530E4496A6PFB3 CRC64;  
 CC SEQUENCE

Query Match 50.6%; Score 875.5; DB 1; Length 351;  
 Best Local Similarity 54.3%; Pred. No. 1.1e-61;  
 Matches 178; Conservative 51; Mismatches 82; Indels 17; Gaps 5;

QY 1 MAMTLLDWCRCNDVNSQRTLLVWGIPIVNCDAEIEETLQAAM-PQVSYRMGLGRMFWREE 59  
 DB 1 MTRLLDWCRCNDVNSQRTLLVWGIPIVNCDAEIEETLQAAM-PQVSYRMGLGRMFWREE 60  
 QY 60 NAKAALLELTGAVDYAAIPREMPGKGVWVKVLPKPTSDAEFLERLHLFLAREGVTVDV 119  
 DB 61 NRKVALVGLTAETSHALVPEKIPGKGIWRVIFKPPDPNTFLSRLNEFLAGEGVTGEL 120  
 QY 120 ARVLGFQNPFTPG-----PEMPAEMLYLDNVIQPLVESIWKRLTLFSGKHGP 170  
 DB 121 TRALAHEN-----GSLDLEQGMIPEMWAPMLAQALE-ALQALQCLYKLVFSGREPP 174  
 QY 171 RAWRGNFDPLWHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLE 230  
 DB 175 EFGEEFGRWFTTQMIKAWQVDPVEKRRRLMESLRGPAADVIRILKSNPAITTVDECL 234  
 QY 231 KALEQVFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAIDKDNVNQAR 290  
 DB 235 QALEEVFGVTNPRELQVKYLTYYQKDEKLSAYVIRLEPLLOKVEKGAIDKDNVNQAR 294  
 QY 291 LEQVIAGANHSAGAIRRQLWLTGAGEGPG 318  
 DB 295 LDQVIAGAVHK-TIRRELNPEDGPAPG 321

## RESULT 6

MOPL\_MOUSE STANDARD; PRT; 352 AA.  
 AC Q9ERH6;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Modulator of apoptosis 1 (MAP-1).  
 GN MOAP1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RC TISSUE=Brain;
RX MEDLINE=21264738; PubMed=11060313;
RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C.,
RA Yu V.C.;
RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
RL associates with Bax through its Bcl-2 homology domains.";
RN J. Biol. Chem. 276:2802-2807(2001).
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6NCr; TISSUE=Hematopoietic, and Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RL libraries.";
RC Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -1- SIMILARITY: Belongs to the PNMA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF305551; AAG31787.1; -.
CC EMBL; AK019599; BAB31810.1; -.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB047632; BAB12156.1; -.
CC EMBL; AB047632; BAB12156.1; -.

```



```

NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain cortex;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Nikaide I., Osato N., Saito R., Saito R., Saito R., Schonbach C., Gojohori T.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konoaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kogawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the PMA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AK043718; BAC31626.1; -
CC EMBL; AK043910; BAC31700.1; -
CC MGD; MGI:2444129; Pma2.
KW Nuclear protein.
FT DOMAIN 333 338 POLY-GLU.
SQ SEQUENCE 365 AA; 41201 MW; 0F90C940B9D843D9 CRC64;
Query Match
Best Local Similarity 45.9%; Score 704; DB 1; Length 365;
Matches 144; Conservative 65; Mismatches 95; Indels 10; Gaps 5;
QY 1 MAMTLLEDWCRGMDVNSORTLLVWGIPVNCDEAETETLQAPQV-SYRMGLRMFWREE 59
Db 1 MAVALLBECKIMGVQKSLVVVDIPVDCGEPIQTVLQELKCVGSYRLGLKIFQKD 60
QY 60 NAKAALLELTGAVDYAIPREMPGKGGVWKVLFKPTSDAEFLRLHLFLAREGWTQDV 119
Db 61 NTSVVLVELMEDTMSVPSEVQKGGVKKVIFKTPQDTEFLQRLNLFLEKSGQTWAGM 120
QY 120 ARVLGFQNPPT-TP---GPENPAEMLNLYLDNVIPQLVSTWYKRLTLFGKGHPRAMRG 175
Db 121 FRALKHGVSPATPCTSPPELLAHLTGQAMVHGQRPPL-PVKYCKMRIFSGSTAAPEEE 179
QY 176 NFDPLAETHNLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECIKALEQ 235
Db 180 PFVWLVQATEIAKEWPIPEAKRWAEBSLRGPAALDMLHIVQADNPNSISVGCLEAFKQ 239
QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQLQKVEKGAIDKDNVNQARLEQVI 295

```

```

Db 240 VFGSTESRSTQVKYLRVYQGEKISAYVLRLETLRRRAVEKRAIPRNIAQVRLQVM 299
QY 296 AGANHSAGAIRQLW 309
Db 300 AGAN-----LGNVLW 309
RESULT 10
YUJK HAEIN
ID YUJK HAEIN STANDARD; PRT; 556 AA.
AC P45127;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ABC transporter ATP-binding protein H11252.
GN H11252.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
RN [2]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-429(2000).
CC -!- SIMILARITY: Belongs to the ABC transporter family. EF3 subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U32805; AAC22902.1; -
CC PIR; G64169; G64169.
CC TIGR; H11252; -
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC Pfam; PF00005; ABC_tran; 2.
CC ProDom; PD000006; ABC_transporter; 2.
CC SMART; SM00382; AAA; 2.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Transport; Repeat; Complete proteome.
FT NP_BIND 40 47 ATP (POTENTIAL).
FT NP_BIND 357 364 ATP (POTENTIAL).
SQ SEQUENCE 556 AA; 62564 MW; FCCB80E23457A493 CRC64;
Query Match
Best Local Similarity 23.6%; Score 104.5; DB 1; Length 556;
Matches 51; Conservative 34; Mismatches 90; Indels 41; Gaps 7;

```

```

QY 90 VLFKEPPTS--DAEFLERLHLFLAREGWTQDVVARVLGFQNPPTTRGPMPAEMLNLYLDN 147
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 185 LLLDEPTNHLDAESVAVLWFLHDYEGTVTAIHD-----RYFLDN 225

QY 148 VTQPLVESYWKRLTLFGSKGHPRAWGNFDPWLEHTNEVLBEQVSDVEKRRRLMESUR 207
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 226 VAGWILE-----LDRGEGIP--WEGNYSWLEQEKRELEQEQATENARQKSTAKED- 274

QY 208 GPAADVIRILKSNPAITTAECCLKALEQVFGSVESRDAQIKFLNTYQNPGEKLSAYVIR 267
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 275 -----EWVRQNPGRQAKSMARFDELNSGYQKRNETNELFI--PPGRRLGDKVIE 326

QY 268 LEPLQKVVEKAIDKDNVQNARLEQV--IAGANHSG 302
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 327 VQNLKSYGDRTLDDLSPSIPKAIVGIIGANGAG 362

RESULT 11
ID KINH SYNRA STANDARD; PRT; 935 AA.
AC O43093;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Kinesin heavy chain (Synkin).
OS Syncephalastrum racemosum.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales;
OC Syncephalastraceae; Syncephalastrum.
OX NCBI_TaxID=13706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98273660; PubMed=9613604;
RA Grunmt M., Pistor S., Lottspeich F., Schliwa M.;
RT "Cloning and functional expression of a 'fast' fungal kinesin.";
RL FEBS Lett. 427:179-84(1998).
CC -!- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT. ITS MOTOR
CC ACTIVITY IS DIRECTED TOWARD THE MICROTUBULE'S PLUS END. THE SPEED
CC OF THIS MOTOR IS 4-5 TIMES FASTER THAN ITS ANIMAL COUNTERPARTS.
CC -!- DOMAIN: Composed of three structural domains: a large globular N-
CC terminal domain which is responsible for the motor activity of
CC kinesin (it hydrolyzes ATP and binds microtubule), a central
CC alpha-helical coiled coil domain that mediates the heavy chain
CC dimerization; and a small globular C-terminal domain which
CC interacts with other proteins (such as the kinesin light chains),
CC vesicles and membranous organelles.
CC -!- SIMILARITY: Belongs to the kinesin-like protein family. Kinesin
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ225894; CAA12647.1; -.
CC PIR; T51930; T51930.
CC DR H5SP; P33176; I8G2.
CC DR InterPro; IPR007420; DUF465.
CC DR InterPro; IPR001752; kinesin_motor.
CC DR Pfam; PF04325; DUF465; 1.
CC DR Pfam; PF00225; kinesin; 1.
CC DR PRINTS; PR00380; KINESINHEAVY.
CC DR SMART; SM00129; KISC; 1.
CC DR PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
CC DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
CC KW Motor protein; Microtubule; ATP-binding; Coiled coil.
CC FT DOMAIN 1 341 KINESIN-MOTOR (BY SIMILARITY).
CC FT DOMAIN 342 887 COILED COIL (POTENTIAL).
CC FT DOMAIN 888 934 GLOBULAR.
CC FT NP_BIND 87 94 ATP (POTENTIAL).

```

```

FT NP_BIND 237 244 ATP (POTENTIAL).
SQ SEQUENCE 935 AA; 105040 MW; CFBDFD33B3827E28 CRC64;

Query Match
Best Local Similarity 6.0%; Score 103; DB 1; Length 935;
Matches 69; Conservative 56; Mismatches 121; Indels 72; Gaps 15;

QY 16 NSQRTLLVWGIPVNCDEAIEETLQAAMPQVSYRMGLGRMFREENAKAALLEL----- 68
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 297 NSRTLLIINCSPSSSNEAETLSTLFGARAKSIKNKAV-----NADLSPAELKALLKKV 351

QY 69 -TGAVDYAAIPREMPGKGVKVLFPKPTSDAEFLERLHLFLAREG-WTVQDVAVLGFQ 126
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 352 KSEAVTYQTYIAALREGVNVRTGTGP-----EGKWTMDKVKSGDFA 395

QY 127 NPTPTPGPMPAEMLNLYLDNVIQPLVESIWKYKLTFLSGKGHPRAWGNFDPWLEHTNE 186
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 396 GLPPAPGFKSP-----VSDEGRPATPV-----PTEKDEREFIKRENE 435

QY 187 VLEEWQVSDVE---KRRRLMESLR---GPAADVIRILKSNPAITT--AECCLKALEQVVF 237
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 436 LMD--QISEKETELTNREKLLESREEMGYKQEQSVTKENQOMTSELSELRLQLQKV- 492

QY 238 GSVESRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKV--VEKAIDKDNVQV--ARLSQ 293
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 493 -SYESKENAIT--VDSLKEANQDLMAELBELKNLSEMQAHKDATDSDEKREKAKFMAQ 549

QY 294 VIAGANHSGAI---RRQL 308
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 550 MMSGFDPGSLINDKERQI 567

RESULT 12
SYEA AERPE STANDARD; PRT; 488 AA.
ID -SYEA AERPE STANDARD; PRT; 488 AA.
AC Q9Y9I6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phenylalanine-tRNA synthetase alpha chain (BC 6.1.1.20)
DE (Phenylalanine-tRNA ligase alpha chain) (PheRS).
DE PHES OR APE2302.
GN Aeropyrum pernix.
OS Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kurohda N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RT DNA Res. 6:83-101(1999).
RL -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC Phe-tRNA synthetase alpha chain subfamily 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----

```





```
DR InterPro; IPR000782; Bigh3_fasciclin.
DR SMART; SM00554; FAS1; 2.
DR PROSITE; PS00213; FAS1; 2.
KW Membrane; Repeat; Signal; Glycoprotein; Lipoprotein; GPI-anchor.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 404 FASCICLIN-LIKE ARABINOGALACTAN PROTEIN 1.
FT PROPEP 405 424 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 25 170 FAS1 1.
FT DOMAIN 184 323 FAS1 2.
FT LIPID 404 404 GPI-anchor amidated glycine (potential).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 424 AA; 44849 MW; 6B256D3CE592881B CRC64;

Query Match
Best Local Similarity 5.7%; Score 98.5; DB 1; Length 424;
Matches 79; Conservative 49; Mismatches 127; Indels 109; Gaps 18;

QY 1 MAMTILDEWCRGMDYNSORTLAVGIPVNCDEAIEE-----TLQAAMPQVSYRMLGRM 54
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
44 LQOTHLAD-----EINRRRTITVCAY---DNAAMSAUTSKGYTLSTLKNILSLHLVDY 94
QY 55 FW-----REBNKAA--LLELTGAVDYAAIPREMPGKGV-----WKVLEKPTSD 98
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
95 FGTKKLHQIRDSALAAATLFOATGAA-----PGTSGFVNITDLRGKKVGFPGDGD 145
QY 99 AEFLERLHLFLARE-----GWTQODVARVGFON-PTTPGPEMPAEMNLVLDNVIQ 151
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
146 -----LSSFFVKSTEEVPYNISIIQISRLVPSETAAPTAA-PAEM---NLTGIMS 193
QY 152 LVESIWYKRLTLFSGKHPRAWGN-----FDPWLEHTNEVLEWQVSDVEKRRRLE 204
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
194 HGCKVFAB--TLITNPGASKTYQESLEGMTVFCPGDDAMKGFGLPKYKMLTAPKKEALD 251
QY 205 SLRGAAPADVIRILKSNPAITTAECLEKALEQVFGSVESRDAQIKPLNTYQNPGEKLSAY 264
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
252 FLAVPTYYSMALKNNGPMNTL-----ATDGANKFELTVQNDGKVTIL- 295
QY 265 VIRLEPLQKVEKGAIDKNNVQARL-----EQVIA-GANHSAGAIRRLMLTGAGRGP 317
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 -----KTRINTVKIVDTLIDEQPLAIYATDKVLLPKELFKASAVEAP 337
QY 318 GPKP 321
Db : : : : :
338 APAP 341
```

```
RESULT 15
YM26 MYCTU
ID YM26 MYCTU STANDARD; PRT; 513 AA.
AC Q10510;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Rv2226/MT2285.
GN Rv2226 OR MT2285 OR MTCY427.07.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean A., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
```

Search completed: September 21, 2004, 13:32:36  
Job time : 10.969 secs

Blank

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 24.8686 Seconds  
(without alignments)  
1272.567 Million cell updates/sec

Title: US-10-037-860-4  
Perfect score: 1729  
Sequence: 1 MAMTLEDWCRGMDVNSORT.....LTGAGEGPGPKPLSVAGADP 329

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	114.5	6.6	2037	2	T16881	hypothetical prote
2	111	6.4	558	2	F81781	ABC transporter At
3	110.5	6.4	559	2	U7667	type II ATP-bindin
4	108	6.2	558	2	H81205	ABC transporter, A
5	104.5	6.0	556	2	G64169	hypothetical prote
6	103	6.0	935	2	T51930	kinesin (imported)
7	101	5.8	473	2	B72457	probable phenylala
8	101	5.8	555	2	C87515	ABC transporter, A
9	98	5.7	513	2	P70776	hypothetical prote
10	98	5.7	554	2	C62743	ABC transporter At
11	97.5	5.6	759	2	B87443	topoisomerase IV,
12	96.5	5.6	564	2	H75607	hypothetical prote
13	95.5	5.5	555	2	H97710	ABC transporter At
14	94.5	5.5	555	2	D82290	ABC transporter, A
15	94.5	5.5	1003	2	C84601	hypothetical prote
16	94	5.4	726	2	E83712	ribonucleoside-dip
17	93	5.4	316	1	C69855	ABC transporter ho
18	93	5.4	659	2	AB0731	alpha-1,2-mannosyl
19	92.5	5.3	464	2	S34024	nonstructural poly
20	92.5	5.3	2431	1	MNWVSF	50S ribosomal prot
21	90.5	5.2	141	2	E81294	hypothetical prote
22	90.5	5.2	671	2	T04457	hypothetical prote
23	90.5	5.2	1027	2	H87316	hypothetical prote
24	90	5.2	513	2	T14864	probable monosacch
25	90	5.2	861	2	G81068	aconitate hydratase
26	90	5.2	1073	2	F89467	protein R09H3.1 [i
27	90	5.2	1660	2	T18501	vitellogenin vit-6
28	89.5	5.2	1140	2	F88349	protein F15D4.7 [i
29	89.5	5.2	1140	2	T20984	hypothetical prote

cytochrome P450 er  
conserved hypothet  
aconitate hydratase  
probable mosk prot  
hypothetical prote  
group II chaperoni  
ryanodine receptor  
transforming prote  
chemotaxis-related  
hypothetical prote  
probable medium-ch  
probable outer mem  
probable leader pe  
chemotaxis protein  
probable transposo  
translation elonga

30 89 5.1 406 1 S18531  
31 89 5.1 555 2 A81074  
32 89 5.1 861 2 A81801  
33 88.5 5.1 330 2 H75353  
34 88.5 5.1 364 2 B83152  
35 88.5 5.1 552 2 JG5617  
36 88.5 5.1 4872 2 S27272  
37 88 5.1 640 1 A55073  
38 88 5.1 1755 2 F82618  
39 87.5 5.1 334 2 S25785  
40 87.5 5.1 559 2 C72732  
41 87.5 5.1 568 2 C72129  
42 87.5 5.1 568 2 G86493  
43 87.5 5.1 764 2 H87322  
44 87.5 5.1 1011 2 E85054  
45 87 461 1 EFH81

ALIGNMENTS

RESULT 1  
T16881  
hypothetical protein T14G12.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C/Accession: T16881  
R/Milicox, L.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans coamid T14G12.  
A:Reference number: Z18596  
A:Accession: T16881  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2037 <MIL>  
A:Cross-references: EMBL:U41268; NID:g1086843; PID:g1086848; PIDN:AAA82437.1; CESP:T14G12.5  
C:Genetics:  
A:Gene: CESP:T14G12.5  
A:Introns: 715/2; 1484/2

Query Match 6.6%; Score 114.5; DB 2; Length 2037;  
Best Local Similarity 22.0%; Pred. No. 1.5;  
Matches 56; Conservative 41; Mismatches 114; Indels 43; Gaps 8;

Qy 36 BETLOAMPQVSYMLGMFWR-EENAKAALLELTGADVAAIIPREMPGKGVNKLFPK 94  
Db 308 EQLLPQAMLAQMTM-GRLLSQLODQARREQTSFMNECREHLRPPAEGSIGQAYSP 366  
Qy 95 PTSDAEFLERLHLFLAREGWTQDVARVLGFQNP-PTRGPEMPAEMLYILDNVLPV 153  
Db 367 DDEGEESQR-----GSSPPVQIPDSRSPGVNFTN----- 400  
Qy 154 ESIWYKRLTLFSGKHPRAWRGNFDPLWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADV 213  
Db 401 ----AKULPKFDGNGFNFRNGFD-----TVVLDPLPSVTCKNLLRNHLVGNQQC 450  
Qy 214 IRLKSNPNPAITAECLKALEQVFGSVESRDRQIKF--LNTYQNPGEKLSAYVIRLEPL 271  
Db 451 IS--HDDPLVAYQTTMDMLESVYKGDQGLERFRKLKFKHQSNEQMKDLTSHQLL 508  
Qy 272 LQKWEKGAIKDKN 285  
Db 509 VQRLVSTGLSATDD 522

RESULT 2

F81781  
ABC transporter ATP-binding protein NMA2101 [imported] - Neisseria meningitidis (strain 2  
C:Species: Neisseria meningitidis  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C/Accession: F81781  
R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.  
A:Reference number: A81775; MUID:20222556; PMID:10761919  
A:Accession: F81781  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-558 <PAR>  
A:Cross-references: GB:AL162758; GB:AL157959; NID:g7380672; PIDN:CAB85317.1; PID:g7380702  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: NMA2101  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 6.4%; Score 111; DB 2; Length 558;  
Best Local Similarity 22.0%; Pred. No. 0.42;  
Matches 51; Conservative 39; Mismatches 92; Indels 50; Gaps 8;

Qy 90 VLFKPEPTS--DAEFERLHLFLAREGTVQDVARVLGFQNPPTPTPGPEPAEMLNYILDN 147  
Db 187 LLLDPTNHLDAESVLEWLSQFLVR-----PFGTVAVTHDRYFLDN 227

Qy 148 VIQPIVESIWYKRLTLTFSKGHPRAGNFDPLWLEHTNEVLVEWQVSVKRRRLMESLR 207  
Db 228 AAEWILE-----LDRGHGIPWKGNYSSWLEQKEKRLNEAKSAARVAKMQBLE 277

Qy 208 -----GPAADVIRILKSNPAITTAECLEKALEQVFGSV-ESSRDAQIKFLNTYQNGP 258  
Db 278 WYRONAKGROAKSKARLAREEENSVEYQKNETQEIFIPVAERLGNIEFVNVSKSFG 337

Qy 259 EKLSAYVIRLEPLQKV---VEKGAIKDKNVQARLEQVITAGANH--SGAIR 305  
Db 338 DK-----VLIDDLFSKVPAGAVIGIIPNGAGKSTLFKMIAGKEQPDGSEVK 384

RESULT 3  
JC7667

type II ATP-binding cassette (ABC) transporter, Mac-1 - *Myxococcus xanthus*  
C:Species: *Myxococcus xanthus*  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
C:Accession: JC7667

R:Kimura, Y.; Yamanishi, Y.; Tokumasu, Y.; Terasaka, H.; Yoshinobu, J.  
J. Biochem. 129, 351-356, 2001  
A:Title: Characterization of the mac-1 gene encoding a putative ABC transporter from *Myxococcus xanthus*  
A:Reference number: JC7667; MUID:21125588; PMID:11226873  
A:Accession: JC7667  
A:Molecule type: DNA  
A:Residues: 1-559 <KIM>  
A:Cross-references: DDBJ:AB041227  
A:Experimental source: ATCC 31046  
C:Comment: This protein contributes to the export or import of a developing signal or a  
C:Genetics:  
A:Gene: mac-1  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
C:Keywords: antibiotic; sporulation

Query Match 6.4%; Score 110.5; DB 2; Length 559;  
Best Local Similarity 21.5%; Pred. No. 0.47;  
Matches 71; Conservative 47; Mismatches 113; Indels 99; Gaps 17;

Qy 60 NAKAALELTGAVDYAAIPREMPCKGGVWVLFKPEPTSDAEPLERLHLFLARE----- 112  
Db 89 NVEGLKEIRATIDRF--NEVSAK-----FAEPMSDAE-MEKL--LAEQGRLODAI 134

Qy 113 ----GW----TVQ-----DVARVLG-----FQNP-----TPTPGPE 135  
Db 135 DAVNGWELDRTIEMAMDALRLPPGDADVTKLSGGKRRVALCRILLEKPDLLLDLEPTNH 194

Qy 136 MPAEML-----NYLDNVIQPLVESIWYKRLTLFSGKHPRAWRG 175  
Db 195 LDAESVAWLQALKEYKGTIVCIITHDRYFLDNAAEWILE-----LDRGEGVP--WKG 244

Qy 176 NFQDFPETHTEVLEEWQVSDVEKRRRLMESLFGPAADVIRILKSNPNPATTTAECLKALEQ 235

A:Note: best homolog was a hypothetical protein from *Escherichia coli*  
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
 F:23-236/Domain: ATP-binding cassette proteins; p-loop  
 F:40-47/Region: nucleotide-binding motif A (P-loop)  
 F:340-518/Domain: ATP-binding cassette homology <ABC2>  
 F:357-364/Region: nucleotide-binding motif A (P-loop)

Query Match 6.0%; Score 104.5; DB 2; Length 556;  
 Best Local Similarity 23.6%; Pred. No. 1.4;  
 Matches 51; Conservative 34; Mismatches 90; Indels 41; Gaps 7;  
 QY 90 VLFPPTS--DAEFLERLHLFLAREGWTVQDVAVLGFQNPPTPGPEMPAEMLNLYLDN 147  
 DB 185 LLLDEPTNHLDAESVAVLERFLHDYEGTVVAITHD-----RYFLDN 225  
 QY 148 VIQPLVESIWKRLTLFSGKHPRAWGNFDPWLEHTNEVLEWQVSVKRRRLMESLR 207  
 DB 226 VAGWILE-----LDRGEGIP--WEGNYSWLEQKEKLEQEQATENARQKSIKEL- 274  
 QY 208 GPAADVIRILKSNPAITTAELCLALEQVFGSVSSRDAQIKELNTYQNGEKL SAYVIR 267  
 DB 275 ----EWRQNPKGQAKKARMAKRDDELNSGEYQKRNETHLFI-----PPGRLGDKVIE 326  
 QY 268 LEPLQKQVEKAIDKDNVNQARLEQV- IAGANHS 302  
 DB 327 VQNLTKSYGDRTLDDLFSIPKGAIVGIGANGAG 362

RESULT 6  
 T51930  
 kinesin [imported] - pin mould (Syncephalastrum racemosum)  
 C:Species: Syncephalastrum racemosum  
 C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 03-Nov-2000  
 C:Accession: T51930  
 R:Gumm, M.; Pistor, S.; Lottspeich, F.; Schliwa, M.  
 FEBS Lett. 427, 79-84, 1998  
 A:Title: Cloning and functional expression of a 'fast' fungal kinesin.  
 A:Reference number: 225870  
 A:Accession: T51930  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-935 <GRU>  
 A:Cross-references: EMBL:AJ225894; PIDN:CAA12647.1  
 C:Superfamily: kinesin heavy chain; kinesin motor domain homology

Query Match 6.0%; Score 103; DB 2; Length 935;  
 Best Local Similarity 21.7%; Pred. No. 4.1;  
 Matches 69; Conservative 56; Mismatches 121; Indels 72; Gaps 15;  
 QY 16 NSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMGLRMFWRENAKAALEL----- 68  
 DB 297 NSRTLLIINCSPSSYNAEILTLRFGARAKSIKNAKV-----NADLSPAELKALLKV 351  
 QY 69 -TGAVDYAALPREMPGKGWVKVLFKPPPTSDAEFLERLHLFLAREG-WTVQDVAVLGFQ 126  
 DB 352 KSEAVTYQTVIAALEGEVNVWRTGCTVP-----EGKWTMDKYSKGDFA 395  
 QY 127 NPTPTGPEMPAEMNLINLVIOPLVESIWKRLTLFSGKHPRAWGNFDPWLEHTNE 186  
 DB 396 GLPPAPGFKSP-----VSDEGSRPATPV-----PTLEKDEREEFIKRENE 435  
 QY 187 VLEWQVSDVE-----KRRRLMESLR---GPAADVIRILKSNPAITT--AECLKALEQVF 237  
 DB 436 LMD--QISEKETELTNREKLESUREMGYYKEQSVTKENQOMTSELSELRLQLOKV- 492  
 QY 238 GSVESSRDAQIKFLNTYQNGEKL SAYVIRLEPLQKV--VEKAIDKNVQ--ARLEQ 293  
 DB 493 -SYESKENAIT--VDSLKEANQDLMAELELKNLSEMRQAHKDATSDSKKAKMAQ 549  
 QY 294 VIAGANHSQAI-----RRQL 308  
 DB 550 MMSGFDPFGILNDEKQI 567

## RESULT 7

B72457

probable phenylalanyl-tRNA synthetase alpha chain APE2302 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: B72457  
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: B72457  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-473 <KAW>  
 A:Cross-references: DBJ:AP000064; NID:G5105945; PIDN:BA081314.1; PID:G5106003  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE2302  
 C:Superfamily: yeast cytosolic phenylalanine-tRNA ligase beta chain

Query Match 5.8%; Score 101; DB 2; Length 473;  
 Best Local Similarity 24.5%; Pred. No. 2.2;  
 Matches 50; Conservative 26; Mismatches 59; Indels 70; Gaps 12;  
 QY 31 DEAE-----IETLQAAMPQVSYRM-----LGRMFWEENAKAALEL----- 68  
 DB 116 DQAEALKTTITPLKLELVNVA SGSKPTVGDLEALSGLRREARRSIVLRKVPAA 175  
 QY 69 --TGAVDYAALPREMPGKGWVKVLFK-----PTSD---AEFLERLHLFLAR 111  
 DB 176 LARARVEAAVLTRDML-KSGEWRELRFKPNVKAEPPLVPARHFLAEFTLERL----- 228  
 QY 112 EGWTVQDVAVLGFQNPPTPGPEMPAEMLNLYL-----DNVIQPLVESIWK--RLTLF 164  
 DB 229 -----RUILRELGPFE---VRGPIVLELNFVDFVLFQADHPAREIHDSLIKSPRRGDL 280  
 QY 165 SGKG-----HPRAWRGNFDP 179  
 DB 281 SGYSDLVERVASVHERGKWRWSP 304

## RESULT 8

C87515

ABC transporter, ATP-binding protein CC2148 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 22-Oct-2001  
 C:Accession: C87515  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonias, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: C87515  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-555 <STO>  
 A:Cross-references: GB:AE005673; NID:G13423641; PIDN:AAK24119.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC2148  
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 5.8%; Score 101; DB 2; Length 555;  
 Best Local Similarity 23.3%; Pred. No. 2.8;  
 Matches 76; Conservative 47; Mismatches 118; Indels 92; Gaps 19;

QY 20 TLLVWG-IPVNCDAEIEETLQAAMPQVSYRMGLRMFWRENAKAALELTLTGAVDYAALP 78  
 DB 84 TLDVWGNVIADCEKQIFDRYNALAAQLGEGYTDLEL--EEMTK--LQEVIDARD----- 134





Db	685	-----SADMDKRRRLHENKMKMLQSPNFHVSRTRLVLIYNDFASWNP-----	720
Qy	230	LKALEQVGVESSRDA-----QIKFLNTYQNPGEKLSAYVIRLEPLLQKVVEKGAI	282
Db	727	-KQLNRLVDAVTSRATKQKPCIRQIKFL-----QNE-----KKGKVD	763
Qy	283	KDNVQ 288	
Db	764	TKNYSR 769	

Search completed: September 21, 2004, 13:40:37  
Job time : 26.8686 secs

```

RESULT 15
C84601
hypothetical protein Ar2g21440 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84601
F:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84601
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1003 <STO>
A:Cross-references: GB:AE002093; NID:g4567275; PIDN:AAD23688.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g21440
A:Map position: 2

Query Match          5.5%; Score 94.5; DB 2; Length 1003;
Best Local Similarity 19.3%; Pred. No. 23;
Matches 59; Conservative 42; Mismatches 74; Indels 131; Gaps 12;

QY   18 QRTLLVWGIFVNCDEABIE-----ETLQAAMPQVSRYRMIGRMFWREENAKAAELLEL 68
      :| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db   560 ERTLFI RNLPFDVTKEEVKQRFTVFGEVESLSVLHKVTKRPGCTAFVKFKTADASVAAI 619
      :| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY   69 TGAVDYAAI PREMPGKG-GYWKVLFKPPTSDAEF-----LERLHLFLARGGWTVQDVA 120
      :| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db   620 SAADTASGVGLLKGRQLNVAVRGVKAADI ELKYTEKNVDHRNLYLAKEG-----672
      :| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY   121 RVLGFPNTPTPGPEMPAEMLN YILDNIQPLVESIWYKRLTLFGSGKHPRAWRGNFDEW 180
      |||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db   673 -----QILDDT--PAEAGV-----684
      |||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY   181 LEHTNEVLEEWSQDVSKRRRL-----MESLRGPAADVIR-----ILKSNNPATITTAEC 229
      :| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:35:22 ; Search time 76.0979 Seconds  
(without alignments)  
1194.261 Million cell updates/sec

Title: US-10-037-860-11

Perfect score: 1462

Sequence: 1 VQKGGVWKVIFKTPNQDTE.....SIEPPEERDGYRWNHGDD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1462	100.0	283	13	US-10-037-860-11
2	755	51.6	149	13	US-10-037-860-9
3	628	43.0	353	9	US-09-965-529-7
4	628	43.0	353	10	US-09-969-680A-7
5	618.5	42.3	463	13	US-10-037-860-13
6	597	40.8	452	16	US-10-408-765A-2385
7	596.5	40.8	351	9	US-09-965-529-1
8	596.5	40.8	351	10	US-09-969-680A-1
9	596.5	40.8	351	12	US-09-804-014A-16
10	596.5	40.8	351	15	US-10-341-434-10
11	593	40.6	195	13	US-10-037-860-7
12	564	38.6	329	13	US-10-037-860-4
13	562.5	38.5	318	12	US-09-804-014A-40
14	560	38.3	321	12	US-09-804-014A-39
15	550	37.6	312	12	US-09-804-014A-73

16	550	37.6	312	12	US-09-804-014A-74	Sequence 74, Appl
17	475.5	32.5	399	15	US-10-094-749-1978	Sequence 1978, Ap
18	394	26.9	403	15	US-10-094-466-38	Sequence 38, Appl
19	378.5	25.9	337	12	US-10-296-115-1208	Sequence 1208, Ap
20	215	14.7	120	12	US-09-804-014A-42	Sequence 42, Appl
21	156.5	10.7	204	14	US-10-029-386-33747	Sequence 33747, A
22	123	8.4	120	12	US-09-804-014A-41	Sequence 41, Appl
23	117	8.0	538	16	US-10-408-765A-2992	Sequence 2992, Ap
24	113	7.7	2383	14	US-10-082-830-260	Sequence 260, App
25	110.5	7.6	584	15	US-10-221-278-355	Sequence 355, App
26	110.5	7.6	584	15	US-10-291-172-355	Sequence 355, App
27	107	7.3	1031	11	US-09-764-875-686	Sequence 686, App
28	107	7.3	1035	15	US-10-158-057-197	Sequence 137, App
29	107	7.3	1459	16	US-10-408-765A-2246	Sequence 2246, Ap
30	106	7.3	542	12	US-10-205-331-57	Sequence 57, Appl
31	105.5	7.2	758	12	US-10-282-122A-67949	Sequence 67949, A
32	102.5	7.0	879	12	US-10-282-122A-60655	Sequence 60655, A
33	102	7.0	116	9	US-09-864-761-34645	Sequence 34645, A
34	102	7.0	225	10	US-09-764-891-4172	Sequence 4172, Ap
35	101.5	6.9	788	14	US-10-128-714-8204	Sequence 8204, Ap
36	100	6.8	750	14	US-10-410-681-12	Sequence 12, Appl
37	100	6.8	1082	16	US-10-437-963-162190	Sequence 162190,
38	99.5	6.8	860	12	US-10-072-012-838	Sequence 838, App
39	99.5	6.8	860	12	US-10-037-417-59	Sequence 59, Appl
40	99.5	6.8	860	15	US-10-080-334-166	Sequence 166, App
41	99	6.8	342	12	US-10-425-114-71718	Sequence 71718, A
42	99	6.8	750	12	US-10-424-599-268662	Sequence 268662,
43	99.5	6.7	520	15	US-10-220-381-12	Sequence 12, Appl
44	98	6.7	1070	14	US-10-420-845-22	Sequence 22, Appl
45	98	6.7	1504	9	US-09-932-145-7	Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-10-037-860-11  
; Sequence 11, Application US/10037860  
; Publication No. US20020123114A1  
; GENERAL INFORMATION:  
; APPLICANT: Jerome B. Posner  
; APPLICANT: Josep O. Dalmau  
; APPLICANT: Myrna R. Rosenfeld  
; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA  
; TITLE OF INVENTION: ANTIBODIES  
; FILE REFERENCE: 2581.1004-004  
; CURRENT APPLICATION NUMBER: US/10/037,860  
; CURRENT FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: 09/189,527  
; PRIOR FILING DATE: 1998-11-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-037-860-11

Query Match	100.0%;	Score 1462;	DB 13;	Length 283;
Best Local Similarity	100.0%;	Pred. No. 9.6e-128;		
Matches 283;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VQKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALQEGVSPATVPCISPEL	60	
Db	1	VQKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALQEGVSPATVPCISPEL	60	
Qy	61	LAHLGQAHAPQPLPMRYKLRVFGSAVPAPBESEFEVWLEQATEIVKWPVTEAE	120	
Db	61	LAHLGQAHAPQPLPMRYKLRVFGSAVPAPBESEFEVWLEQATEIVKWPVTEAE	120	
Qy	121	KRWLAESLRGPALDLMHIVQADNPISVSECELEAFQVFGSLESRTTAQVRYLKTQEE	180	
Db	121	KRWLAESLRGPALDLMHIVQADNPISVSECELEAFQVFGSLESRTTAQVRYLKTQEE	180	



```
Db 139 EMLNVLNDVITQPLVESIWYKRLTLFSGKDIIPGCEETFDPMLEHTNVELEWQVSDVEK 198
Qy 122 KRWLAESLRGALDLMHIVQADNPISVVEECLEAFKQVFGSLERSTTAQVRYLKYQEG 181
Db 199 RRRLMESURGAADVIRILKNNPAITTAECILKALEQVFGSVESRDQAIKFLNTYQNG 258
Qy 182 EKVSAVYLRLETLRKAVEKRAIPRIADQVRLEQVMAGA---TLNQMLWCLRELKQD 237
Db 259 EKL SAYVIRLEPLQKVKEKADKNVYNQARLEQVIAGAHSGAIRRLQWL---TGAGE 315
Qy 238 GPPFSFLELMKVIRREEREEASFENES 264
Db 316 GPAPNLFQLLVQIREEEAKKEEEAE 342
```

## RESULT 5

```
US-10-037-860-13
; Sequence 13, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrta R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; FILE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 463
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-13
```

```
Query Match 42.3%; Score 618.5; DB 13; Length 463;
Best Local Similarity 50.2%; Pred. No. 8.3e-49;
Matches 135; Conservative 44; Mismatches 85; Indels 5; Gaps 3;

Qy 1 VQKGGVWKVIFKTPNQDTFELRLNLFLEKEGQTVSGMFRALGOEGVSPATVPCISPEL 60
Db 82 IPKGGPWEVIVKRNDSGEFLNLFLEERITVSDMRVIGSDTNCSPRTVTSPEF 141
Qy 61 LAHLGQMAHAPOPLL-PMYRKLRVFGSAVPAPEESFEVWLEQATEIVKEWPVTEA 119
Db 142 WT--WAQTLGAAVQPLLEOMLYRELRFVSGNTISIPGALAFDAWLEHTMQLQWQVPEG 199
Qy 120 EKRWLAESLRGALDLMHIVQADNPISVVEECLEAFKQVFGSLERSTTAQVRYLKYOE 179
Db 200 EKRRLMELRGALQVQVSGLRASASITVEECLEALQVQVPGVESHKIAQVLCRAYQE 259
Qy 180 EKVSAVYLRLETLRKAVEKRAIPRIADQVRLEQVMAGATLQMLWCLRELKQDGP 239
Db 260 AGEKVSFVLRLEPLLQRAVENNVSRNNVNTFLKEVLSGATLPDKLRDKLMLKQRRK 319
Qy 240 PPSFLELMKVIRREEREEASF--ENESIE 266
Db 320 PPGFLALVKLLREEEWEATLGPDPRESLE 348
```

## RESULT 6

```
US-10-408-765A-2385
; Sequence 2385, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojn D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
```

```
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2385
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2385
```

```
Query Match 40.8%; Score 597; DB 16; Length 452;
Best Local Similarity 46.3%; Pred. No. 8.1e-47;
Matches 132; Conservative 48; Mismatches 79; Indels 26; Gaps 5;

Qy 1 VQKGGVWKVIFKTPNQDTFELRLNLFLEKEGQTVSGMFRALG-----QEGVSPATVPC 55
Db 86 IPKGGGWEVIVKRNPDDEFSLRNLNLFLEKDEGRSMTDVARALGCCSLPAESLDAEVMPQ 145
Qy 56 I-SPELLAHLGQMAHAPOPLLPMYRKLRVFGSAVPAPEESFEVWLEQATEIVKEW 114
Db 146 VRSPL-----EPPKESMWYRKLVFGSGTASPSGEETFDQVTEIMPIW 193
Qy 115 PVTEAEKRWLAESLRGALDLMHIVQADNPISVVEECLEAFKQVFGSLERSTTAQVRYL 174
Db 194 QVSEVEKRRRLLESRLGPALESIMVQLQANDSITVEQCLDALKQIFGDKEDFRASQRFPL 253
Qy 175 KTYOEKEKVSAYVIRLETLRKAVEKRAIPRIADQVRLEQVMAGATLQMLWCLREL 234
Db 254 QTSKIGEKVSTFLRLLEPLLQKAVHKSPLSVRSTDMIRLKHILARVAMTALRGKLEL 313
Qy 235 KDQGPSPSFLLELMKVIRREEE---EEASFENESIEEPEERDGYGR 276
Db 314 DQRCPPNFLEMLKLIIRDEEWENTEAVMKNK-----EKPSGRGR 353
```

## RESULT 7

```
US-09-965-529-1
; Sequence 1, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 112301CD1
US-09-965-529-1
```

Query Match 40.8%; Score 596.5; DB 9; Length 351;

		Best Local Similarity	48.3%;	Pred. No.	6.3e-47;	Gaps	5;
		Matches	128;	Conservative	49;	Mismatches	75;
		Indels	13;				
Qy	1	VQGKGWVKVIFKTPNODTFLERLNLFLEKEGOTVSGMFRALQEGVSPATVPCISPEL	60				
Dd	82	IPQGGGIWRVIRPDPDNPTFLSRLNFLAGEGMTUGELSRALCHENGSLDPEGMLPEM	141				
Qy	61	LARLLGQAHAHAPQLLP-MRYRKLRIYFGSAVPAAPEESFEVWLQEATEIVIKEWNPYTEA	119				
Dd	142	WAPMLAQAL-EALQPALQCLKYKKLRVFSGRESPEPGEEFFGRWMFTTMQIKAWQPDV	200				
Qy	120	EKRRLAESRGPALDLMHIVOADNPISVEECLEAKQVFGSLESRTTAQVRYLKTYYOE	179				
Dd	201	EKRRLLESRGPAIDVIRVLKINPLITVDCEIQALEEVEVGVTNDPRELOKVLYTYQK	260				
Qy	180	EGEKSVAIVLRLETILRKAVEKRAIPRIAQVRLEQVMAGA---TLNQMLWCRLRELKD	236				
Dd	261	DEEKLSAYVLRLEPLLQKLVORGAIERDAVNQAOLDQVIAGVHKTIIRREL-----NLPE	315				
Qy	237	QGPPPSFLELMKWIRE---EEEEEA	258				
Dd	316	DGPAGFLOLIIVLIKOVEAAEEEEEA	340				

```

RESULT 8
US-09-969-680A-1
; Sequence 1, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/149,641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164,203
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030124649A1 112301CD1
US-09-969-680A-1

```

	Query Match	40.8%;	Score 596.5;	DB 10;	Length 351;
	Best Local Similarity	48.3%;	Pred. No. 6.3e-47;		
	Matches 128;	Conservative 49;	Mismatches 75;	Indels 13;	Gaps 5
Qy	1	VQGGVWKVFKTPNQDTEFLERLNLFLEKGGQTVSGMFRALGQGVSPATVPCISPEL	60		
Db	82	IPGKGIVRVFKPPDPDNTFLSRINEFLAGSGMTVGLSRALGHENGSLDPEQGMIPEM	141		
Qy	51	LAHLIGQAMAHAPQLLP-MRYRKLKRVFSGSVAPEERESFVWLEAQATEIKVENPVTGA	119		
Db	142	WAPMLAQAL-EALQALQCLKTKLVFSGRSPSPGEEEFGRWFPHITQMIKAWQVPDV	200		
Qy	120	EKKRWLAESLRGALDLMHIVQADNPISIVBECLEAFKQVFGSLESRRTAQVRYLKTQYE	179		
Db	201	EKKRRLLSRLGFALDIVRLKINPNLLITVDECLQALEEVFGVTDNPRLOVKYLLTYQK	260		
Qy	180	EKEKVSAYVLRLETLRLKAVEKRAIPRRITADQVRLQEQVMAGA---TLNQMLMCRLRELKD	236		

```

DB      261 DEEKL SAYVLRLPLELQKLVORCAIERDAVNQARLDQVIAGAVHKIRREL-----NLPE 311
OY      237 QGPPSPSFLMLMKVIRE-----EEEEEA 258
        ||| :||: :|||
DB      316 DGPAPGFLLVLILIKDYAEAEEEEA 340

```

RESULT 9  
US-09-804-014A-16  
; Sequence 16, Application US/09804014A  
; Publication NO. US20030064489A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Li  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Vernet, Corine  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Majumder, Kumud  
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-721 US  
; CURRENT APPLICATION NUMBER: US/09/804,014A

```

; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-16

Query Match      40.8%; Score 596.5; DB 12; Length 351;
Best Local Similarity 48.3%; Pred. NO. 6.3e-47;
Matches 128; Conservative 49; Mismatches 75; Indels 13; Gaps 5

Qy      1  VQGGGYYKVF1FKPNQDTEFLRLNLFLEKGGQTVSGMFRALGQGVGSPATVPCISPEL 60
Db      82  IPGGGIVKRV1FKPPDPDNTFLSLNEFLAGEGTVGLSRALGHENGSLDPEQGMIPDM 141

Qy      61  LAHLGQAMAHAPQPLP-MRYKRLRVFSGSAVPAPESESEFVWLEQATEIVKEWPVTEA 119
Db      142  WAPMLAQAL-EALQAPALQCLYKKLRVFSGRESPEFGEEFGFWMFHTTQM1KAWQVPDV 200

Qy      120  EKKRWLAESIRGPALDMHIVQADNPSTSVSECLFAFKQVFGSLESRRTAQVRYLKTQYE 179
Db      201  EKRRLLGESLRGPAIDVIRVUKINNPFLITVDECLQALEEVFGVTNPRLQVKLYLTYYOK 260

Qy      180  EGEKVSAYVLRLETLRLKXAVEKRAIPRR1ADQVRLEQVWVGA---TLNQMLMCRURELKD 236
Db      261  DEEKL SAYVLRLEPLQLKLVQRGATERDAVNGQARLDQVIAGAVHKTRREL-----NLPE 315

Qy      237  QGPPSPFLELMKVIRE---EEEEEA 258
Db      316  DGPAPGFLQLLVLIKDYEAABEEEEA 340

RESULT 10
US-10-341-43A-10
; Sequence 10, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-10

Query Match      40.8%; Score 596.5; DB 15; Length 351;
Best Local Similarity 48.3%; Pred. No. 6.3e-47;
Matches 128; Conservative 49; Mismatches 75; Indels 13; Gaps 5;

Qy      1  VQGGGVKVIKFTPNQDTEFLRLNLFLEKEGQTVSGMFRALGQGVSPATVPCISP 60
Db      82  IPGGGWRVLFKPPDPNTFLSLNEFLAGEGTVGSLRALGHENGSLDPBQGMIP 141

Qy      61  LAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPPEESFEVWLEQATEIVKEMPVTEA 119
Db      142  WAPMLAQAL-EALQAPQCLYKLRVFSGRESPEPGEEFRGWFHTTQMIKAWQVDPV 200

Qy      120  EKKWLAEISLGPALDMHIVQADNPISVBECEAFKQVFGSLESRRTAQVRYLKYQE 179
Db      201  EKRRLESLSLGPALDVIRVLKINPLITVDECLQALEEVEGVTDNPRELQVKYLTIVQK 260

Qy      180  EGEKVSAYVLRLETLRKAVEKRAIPRIADQVRLQVWAGA---TLNQMLWCLRELKD 236
Db      261  DEKLSAYVLRLETLRKAVEKRAIPRIADQVRLQVWAGA---TLNQMLWCLRELKD 236

Qy      237  QGPPPSFLELMKVIRE---EEEEEA 258
Db      316  DGPAPGFLQLLVLIKDYEAEEEEEA 340

RESULT 11
US-10-037-860-7
; Sequence 7, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-7

Query Match      40.6%; Score 593; DB 13; Length 195;
Best Local Similarity 38.3%; Pred. No. 5.9e-47;
Matches 113; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  VQGGGVKVIKFTPNQDTEFLRLNLFLEKEGQTVSGMFRALGQGVSPATVPCISP 60
Db      81  VQGGGVKVIKFTPNQDTEFLRLNLFLEKEGQTVSGMFRALGQGVSPATVPCISP 140

```

```

Qy      61  LAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPPEESFEVWLEQATEIVKEMP 115
Db      141  LAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPPEESFEVWLEQATEIVKEMP 195

RESULT 12
US-10-037-860-4
; Sequence 4, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-4

Query Match      38.6%; Score 564; DB 13; Length 329;
Best Local Similarity 47.0%; Pred. No. 6.2e-44;
Matches 117; Conservative 47; Mismatches 73; Indels 12; Gaps 4;

Qy      3  GKGGVKVIKFTPNQDTEFLRLNLFLEKEGQTVSGMFRALGQGVSPATVPCISP 62
Db      83  GKGGVKVIKFTPNQDTEFLRLNLFLEKEGQTVSGMFRALGQGVSPATVPCISP 138

Qy      63  HLLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPPEESFEVWLEQATEIVKEMPVTEAK 121
Db      139  EMLNYILDNVIQPLVESIWKRLTLFGKGPRAWRGDFDPLWHTNEVLESEWQVSDVEX 198

Qy      122  KRLAESLRGPAALDMHIVQADNPISVBECEAFKQVFGSLESRRTAQVRYLKYQE 181
Db      199  RRLMESLRGPAADVIRILKSNPFAITTAECLEKALEQVFGSVESRRDAQIKFLNTYQNP 258

Qy      182  EKSAYVLRLETLRKAVEKRAIPRIADQVRLQVWAGA---TLNQMLWCLRELKDQ 237
Db      259  EKLAYSIVLRLETLRKAVEKRAIPRIADQVRLQVWAGA---TLNQMLWCLRELKDQ 315

Qy      238  GPPPSFLEL 246
Db      316  GPGPKPLSV 324

RESULT 13
US-09-804-014A-40
; Sequence 40, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10

```

;; PRIOR APPLICATION NUMBER: 60/189,139  
;; PRIOR FILING DATE: 2000-03-14  
;; PRIOR APPLICATION NUMBER: 60/189,140  
;; PRIOR FILING DATE: 2000-03-14  
;; PRIOR APPLICATION NUMBER: 60/190,401  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: 60/190,231  
;; PRIOR FILING DATE: 2000-03-17  
;; NUMBER OF SEQ ID NOS: 75  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 40  
;; LENGTH: 318  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; NAME/KEY: VARIANT  
;; LOCATION: (20)  
;; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the  
;; OTHER INFORMATION: specification  
US-09-804-014A-40

Query Match 38.5%; Score 562.5; DB 12; Length 318;  
Best Local Similarity 48.9%; Pred. No. 8.1e-44;  
Matches 113; Conservative 45; Mismatches 64; Indels 9; Gaps 3;

QY 3 GGGGKGVKVIKFTPNQDTEFLERLNLFLEKEGQTVSGMFRALGOEGVSPATVPCISPELLA 62  
Db 83 GGGGKGVKVIKFTPNQDTEFLERLNLFLEKEGQTVSGMFRALGOEGVSPATVPCISPELLA 138  
QY 63 HLLGQAMAHAPQLLP-PMRYRKLRFVSGSAVPAPEESESFEVWLEQATEIVKWPVTEAK 121  
Db 139 EMLNVLNDVQLPVSIVYKRLTFSGKHPRAMRGDFDPWLEHTNEVLEWQVSDVEK 198  
QY 122 KRWLASLRGPDALDMLHIVQADNPSISVBECEAFQVFGSLSRRTAQVRVLYKTYOEG 181  
Db 199 RRRLESLRGPDALDMLHIVQADNPSISVBECEAFQVFGSLSRRTAQVRVLYKTYOEG 258  
QY 182 EKVSAYVLRLETLRKAVKRAIPRRIADQVRLEQVMAGA---TLNOMLW 228  
Db 259 EKLSAYVLRLETLRKAVKRAIPRRIADQVRLEQVMAGA---TLNOMLW 309

RESULT 14  
US-09-804-014A-39  
; Sequence 39, Application US/09804014A  
; Publication No. US20030064489A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Li  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Vernet, Corine  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Majumder, Kumud  
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-721 US  
; CURRENT APPLICATION NUMBER: US/09/804,014A  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: 60/188,316  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 60/188,277  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 60/189,139  
; PRIOR FILING DATE: 2000-03-14  
; PRIOR APPLICATION NUMBER: 60/189,140  
; PRIOR FILING DATE: 2000-03-14  
; PRIOR APPLICATION NUMBER: 60/190,401  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/190,231  
; PRIOR FILING DATE: 2000-03-17  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-804-014A-73

Query Match 37.6%; Score 550; DB 12; Length 312;  
Best Local Similarity 50.9%; Pred. No. 1.2e-42;  
Matches 113; Conservative 41; Mismatches 66; Indels 2; Gaps 2;

QY 1 VQGGKGVKVIKFTPNQDTEFLERLNLFLEKEGQTVSGMFRALGOEGVSPATVPCISPELL 60  
Db 82 IPGGGIWRVIFKPPDPDNTFLSLNEFLAGEGTVGELSRALGHENGLDPEQGMIPDM 141

;; LENGTH: 321  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-804-014A-39

Query Match 38.3%; Score 560; DB 12; Length 321;  
Best Local Similarity 48.6%; Pred. No. 1.4e-43;  
Matches 119; Conservative 44; Mismatches 72; Indels 10; Gaps 4;

QY 1 VQGGKGVKVIKFTPNQDTEFLERLNLFLEKEGQTVSGMFRALGOEGVSPATVPCISPELL 60  
Db 82 IPGGGIWRVIFKPPDPDNTFLSLNEFLAGEGTVGELSRALGHENGLDPEQGMIPDM 141  
QY 61 LAHLGQAMAHAPQLLP-PMRYRKLRFVSGSAVPAPEESESFEVWLEQATEIVKWPVTEA 119  
Db 142 WAPMLAQAL-EALQPALQCLVKYKLRVFSRGSEPEEGEEFGRMFHTTQMILKAWQVDPV 200  
QY 120 EKKRWLASLRGPDALDMLHIVQADNPSISVBECEAFQVFGSLSRRTAQVRVLYKTYOEG 179  
Db 201 EKKRWLASLRGPDALDMLHIVQADNPSISVBECEAFQVFGSLSRRTAQVRVLYKTYOEG 260  
QY 180 EGEKVSAYVLRLETLRKAVKRAIPRRIADQVRLEQVMAGA---TLNOMLW 236  
Db 261 DEKLSAYVLRLETLRKAVKRAIPRRIADQVRLEQVMAGA---TLNOMLW 315  
QY 237 QGPPPP 241  
Db 316 DGAPAP 320

RESULT 15  
US-09-804-014A-73  
; Sequence 73, Application US/09804014A  
; Publication No. US20030064489A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Li  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Vernet, Corine  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Majumder, Kumud  
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-721 US  
; CURRENT APPLICATION NUMBER: US/09/804,014A  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: 60/188,316  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 60/188,277  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 60/189,139  
; PRIOR FILING DATE: 2000-03-14  
; PRIOR APPLICATION NUMBER: 60/189,140  
; PRIOR FILING DATE: 2000-03-14  
; PRIOR APPLICATION NUMBER: 60/190,401  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/190,231  
; PRIOR FILING DATE: 2000-03-17  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 73  
; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-804-014A-73



B/ank

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 21.0409 Seconds  
(without alignments)  
694.369 Million cell updates/sec

Title: US-10-037-860-11  
Perfect score: 1462  
Sequence: 1 VQKGGVWVIFKTPNQDTE.....SIEEPERDGYGRWNHGGD 283

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pap.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pap.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pap.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pap.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pap.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	618.5	42.3	462	4	US-09-189-527-13
2	593	40.6	195	4	US-09-189-527-7
3	564	38.6	329	4	US-09-189-527-4
4	100	6.8	750	4	US-09-585-173B-12
5	98	6.7	1070	3	US-08-922-635-22
6	98	6.7	1504	4	US-09-364-206-2
7	97.5	6.7	651	3	US-08-650-766-6
8	97.5	6.7	651	3	US-08-922-635-5
9	97.5	6.7	651	4	US-09-389-487-6
10	95	6.5	754	4	US-09-585-173B-51
11	94	6.4	1560	4	US-09-264-512B-2
12	93.5	6.4	331	3	US-08-556-419-25
13	93	6.4	1805	1	US-07-833-313-2
14	92	6.3	736	4	US-08-252-991A-19048
15	92	6.3	1898	1	US-08-056-200-94
16	92	6.3	1898	2	US-08-800-644-94
17	91.5	6.3	300	4	US-09-252-991A-23947
18	91	6.2	497	4	US-09-345-473E-8
19	91	6.2	518	3	US-09-329-418-3
20	91	6.2	518	3	US-09-329-418-4
21	91	6.2	518	3	US-09-329-418-5
22	91	6.2	518	3	US-09-329-418-9
23	91	6.2	518	3	US-09-531-914-3
24	91	6.2	518	3	US-09-531-914-4
25	91	6.2	518	3	US-09-531-914-5
26	91	6.2	518	3	US-09-531-914-9
27	91	6.2	1786	3	US-08-973-462-8

28 90.5 6.2 420 3 US-09-329-418-8 Sequence 8, Appli  
29 90.5 6.2 420 3 US-09-531-914-8 Sequence 8, Appli  
30 90 6.2 257 4 US-09-107-532A-6287 Sequence 6287, Ap  
31 90 6.2 592 2 US-08-736-770-6 Sequence 6, Appli  
32 90 6.2 592 4 US-09-702-705-1809 Sequence 1809, Ap  
33 90 6.2 592 4 US-09-736-457-1809 Sequence 1809, Ap  
34 90 6.2 592 4 US-09-643-657-4 Sequence 4, Appli  
35 90 6.2 592 4 US-09-671-325-1809 Sequence 1809, Ap  
36 89.5 6.1 800 4 US-09-555-790A-2 Sequence 2, Appli  
37 89 6.1 344 1 US-08-843-993-3 Sequence 3, Appli  
38 89 6.1 344 3 US-09-059-520A-3 Sequence 3, Appli  
39 89 6.1 344 3 US-09-334-275-3 Sequence 3, Appli  
40 89 6.1 1581 4 US-09-866-108A-15754 Sequence 15754, A  
41 88.5 6.1 182 2 US-08-602-941-2 Sequence 2, Appli  
42 88.5 6.1 182 3 US-09-961-264-2 Sequence 2, Appli  
43 88.5 6.1 182 4 US-09-442-099A-2 Sequence 2, Appli  
44 88.5 6.1 182 4 US-09-442-099A-17 Sequence 17, Appli  
45 88.5 6.1 182 4 US-09-612-342-2 Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-09-189-527-13  
; Sequence 13, Application US/09189527A  
; Patent No. 6387639  
; GENERAL INFORMATION:  
; APPLICANT: Jerome B. Posner  
; APPLICANT: Josep O. Dalmau  
; APPLICANT: Myrna R. Rosenfeld  
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma  
; TITLE OF INVENTION: Antibodies  
; FILE REFERENCE: SLK98-01  
; CURRENT APPLICATION NUMBER: US/09/189,527A  
; CURRENT FILING DATE: 1998-11-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13  
; LENGTH: 462  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-189-527-13

Query Match 42.3%; Score 618.5; DB 4; Length 462;  
Best Local Similarity 50.2%; Pred. No. 1.9e-56;  
Matches 135; Conservative 44; Mismatches 85; Indels 5; Gaps 3;  
QY 1 VQKGGVWVIFKTPNQDTEFLERLNLFLKEGQTUSGMPALQGEVSPATVPCISPEL 60  
Db 76 IPKGGVWVIVPRNSDGEFLNRLNRFLEERTVSDMNRVLGSDTNCSPRTVISPEF 135  
QY 61 LAHLLGAMAHAPQLL-PMRYKLRVFGSVAFPAPPEESFEVWLEQATEIVKEMPVTEA 119  
Db 136 WT--WAQTLCGAQVPLLEQMLYRELRFVSGNTISIPGALAFDAWLEHTTLMQWQPEG 193  
QY 120 EKRWLAESIRGPAIDLMHIVQADNPISVVECLEAFQVFGSLESRRTAQVRLKTYQE 179  
Db 194 EKRRIMECLRGPAIQVSGLRASNASITVEECALAQVFGPVGSHKIAQVCLKRAYQE 253  
QY 180 EGEKVSAYVLRLETLARKAVEKAIPIRRADQVRLQVQVAGATINQMLWCRRLKDKQGP 239  
Db 254 AGEKVSFVLRLEPLQRAVENNVSRNNVQTRLRKRVLSGATLPDKLRDKLMKQRRK 313  
QY 240 PPSFLEIMKVIREEEEESAF--ENESIE 266  
Db 314 PPGFLALVLLKREEEWEATLGPDRSLE 342

RESULT 2  
US-09-189-527-7  
; Sequence 7, Application US/09189527A  
; Patent No. 6387639

```

; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-7

Query Match          40.6%; Score 593; DB 4; Length 195;
Best Local Similarity 98.3%; Pred. No. 2,5e-54;
Matches 113; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VQKGGVWVIFKTPNQDTFELRLNLFLEKGGQTVSGMFRALQEGVSPATVPCISP 60
Db 81 VQKGGVWVIFKTPNQDTFELRLNLFLEKGGQTVSGMFRALQEGVSPATVPCISP 140
QY 61 LAHLGQAMAHAPQPLPMRYKLRVFGSAVPAPEESFEVWLEQATEIVKEWP 115
Db 141 LAHLGQAMAHAPQPLPMRYKLRVFGSAVPAPEESFEVWLEQATEIVKEWP 195

RESULT 3
US-09-189-527-4
; Sequence 4, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-4

Query Match          38.6%; Score 564; DB 4; Length 329;
Best Local Similarity 47.0%; Pred. No. 6.2e-51;
Matches 117; Conservative 47; Mismatches 73; Indels 12; Gaps 4;

QY 3 GKGGVWVIFKTPNQDTFELRLNLFLEKGGQTVSGMFRALQEGVSPATVPCISP 62
Db 83 GKGGVWVIFKTPNQDTFELRLNLFLEKGGQTVSGMFRALQEGVSPATVPCISP 138
QY 63 HLHGQAMAHAPQPLPMRYKLRVFGSAVPAPEESFEVWLEQATEIVKEWPVTEAK 121
Db 139 EMLNYILDNIQPLVESIWYKRTLFSGKHPRAWGNFDPWLEHTNEVLEEWQVSDVEK 198
QY 122 KRLAESLRGALDLMHIVQADNPSISVECELEAFKQVFGSLERRTAQVRLKTYCEG 181
Db 199 RRLMESLRGPAADVIRILKSNPAITAECLKALEQVFGSVSSDAQIKLNTYQNGP 258
QY 182 EKSAYVRLLETLTKAVEKRAIPRIADQVRLQVMAAGATINQMLWCLRELKQ 237
Db 259 EKSAYVIRLEPLQKVEKAIDKDNVNQARLEQVIGANHSARIRQLWL---TCAGE 315
QY 238 GPPPSFLEL 246
; TYPE: PRT
; LENGTH: 1070
; TYPE: PRT

```

```

Db 316 GPGPKPLSV 324

RESULT 4
US-09-585-173B-12
; Sequence 12, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585,173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Glycine max
US-09-585-173B-12

Query Match          6.8%; Score 100; DB 4; Length 750;
Best Local Similarity 23.0%; Pred. No. 0.14;
Matches 64; Conservative 40; Mismatches 94; Indels 80; Gaps 13;

QY 21 FLERLNLFEKEG-----QTVSGMFRALQEGVS-----PATVPCISP-----LLAHL 65
Db 214 YVDEINLL--DEGISNLLNLVSEGVNTVEREGISFKHPCRPILLIATYNPEGAVREHLL 271
QY 66 GQAMAHAPQPLPMRYKLRVFGSAVPAPEES--FEVWLEQ-----ATEIVKE 113
Db 272 DRIAINLSAD-LPMSFENRVAAGVATEQENSSQVFWEETDNAKTQIILAREYLKD 330
QY 114 WPVTEAEKKRWLAESLRGALDLMHIVQADNPSISVECELEAFKQVFGSLERRTAQVRY 173
Db 331 VILNRDQKLVVIEALRGCGQG--H--RAELFAARVAKCLAA-----LEGRE----- 373
QY 174 LKTYQEEGKVSAYVRLLETLTKAVEKRAIPRIADQVRLQVMAAGATINQMLWCLREL 233
Db 374 -KVVYDD-----LKKAVELVILPSIITESPPDQ----- 401
QY 234 LKDGQPPSPFLELMKVIREEEESAEFENESIEPEER 271
Db 402 -QNQPPPPPPPPPPQNGESGEEQNEEEDDDKDENEEQ 438

RESULT 5
US-08-922-635-22
; Sequence 22, Application US/08922635A
; Patent No. 6033871
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
; TITLE OF INVENTION: AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6033871
; CURRENT APPLICATION NUMBER: US/08/922,635A
; CURRENT FILING DATE: 1997-09-03
; EARLIER APPLICATION NUMBER: 08/650,766
; EARLIER FILING DATE: 1996-05-20
; EARLIER APPLICATION NUMBER: 60/012,600
; EARLIER FILING DATE: 1996-03-01
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1070
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-08-922-635-22

Query Match          6.7%; Score 98; DB 3; Length 1070;
Best Local Similarity 22.6%; Pred. No. 0.39;
Matches 65; Conservative 33; Mismatches 103; Indels 86; Gaps 14;

QY 30 EKEGQTVSGMFRAL-----GQGVSPATVPCISPELLAHLGQAMAHAP 73
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 10 EKELDTVE-VLKATQKAKEVSKLSNPEKKGDSRLSAAFCIRPSSSPPTVAPASALP 68
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 74 QPLLPMYRKLRFVSGSAVPAPPEESFEVWLEQATEIVKE-WPVTEAEKKRWLAESLRG- 131
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 69 QPIL-----SNQIMFVQEEALASSLSSTDSLTPHQPIAQQ-----CSDSLESI 113
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 132 PA-----LDLMIHQADNPSISVEECLEAFKQVF---GSLSRRTAQVRYLKYQEE 180
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 114 PAGQAASDDLDRDVPAGVGASP-----EHAPEVQVVPFGSGOIIFLPFTCIGYTATNQD- 167
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLQVQMAGATLNQMLWCLRELKDKQPPP 240
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 168 -----FIQLSTLIRQAIE-RQLP-----AWIEAANQREEGQG 199
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 241 PSFLELMKVIREEEERASFEENESIE-----EPEERDGYGRWNHGGDD 283
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 200 EQGEE---EDEEEEDVAENRYFEMGPPDVEEEGGGQGEEREE 243
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

```

```

RESULT 6
US-09-364-206-2
; Sequence 2, Application US/09364206
; Patent No. 6475752
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Baugh, Mariah R.
; TITLE OF INVENTION: Mammalian Imidazoline Receptor
; FILE REFERENCE: PC-0006 US
; CURRENT APPLICATION NUMBER: US/09/364,206
; CURRENT FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY:
; OTHER INFORMATION: 129581CD1
; PUBLICATION INFORMATION:
US-09-364-206-2

```

```

Query Match          6.7%; Score 98; DB 4; Length 1504;
Best Local Similarity 22.6%; Pred. No. 0.66;
Matches 65; Conservative 33; Mismatches 103; Indels 86; Gaps 14;

QY 30 EKEGQTVSGMFRAL-----GQGVSPATVPCISPELLAHLGQAMAHAP 73
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 444 EKELDTVE-VLKATQKAKEVSKLSNPEKKGDSRLSAAFCIRPSSSPPTVAPASALP 502
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 74 QPLLPMYRKLRFVSGSAVPAPPEESFEVWLEQATEIVKE-WPVTEAEKKRWLAESLRG- 131
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 503 QPIL-----SNQIMFVQEEALASSLSSTDSLTPHQPIAQQ-----CSDSLESI 547
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 132 PA-----LDLMIHQADNPSISVEECLEAFKQVF---GSLSRRTAQVRYLKYQEE 180
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 548 PAGQAASDDLDRDVPAGVGASP-----EHAPEVQVVPFGSGOIIFLPFTCIGYTATNQD- 601
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLQVQMAGATLNQMLWCLRELKDKQPPP 240
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 602 -----FIQLSTLIRQAIE-RQLP-----AWIEAANQREEGQG 633
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

```

```

QY 241 PSFLELMKVIREEEERASFEENESIE-----EPEERDGYGRWNHGGDD 283
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 634 EQGEE---EDEEEEDVAENRYFEMGPPDVEEEGGGQGEEREE 677
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

RESULT 7
US-08-650-766-6
; Sequence 6, Application US/08650766D
; Patent No. 6015690
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A HUMAN IMIDAZOLINE RECEPTOR AND
; TITLE OF INVENTION: METHOD FOR CLONING THE SAME
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6015690
; CURRENT APPLICATION NUMBER: US/08/650,766D
; CURRENT FILING DATE: 1996-05-20
; EARLIER APPLICATION NUMBER: US 60/012,600
; EARLIER FILING DATE: 1996-03-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-650-766-6

```

```

Query Match          6.7%; Score 97.5; DB 3; Length 651;
Best Local Similarity 23.0%; Pred. No. 0.21;
Matches 59; Conservative 30; Mismatches 99; Indels 69; Gaps 12;

QY 44 GQGVSPATVPCISPELLAHLGQAMAHAPQPLLPMYRKLRFVSGSAVPAPPEESFEVW 103
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 61 GGEDSLQSAAPCIRPSSSPPTVAPASASLPQIL-----SNQIMFVQEEALASS 110
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 104 LEQATEIVKE-WPVTEAEKKRWLAESLRG-PA-----LDLMIHQADNPSISVEECL 153
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 111 LSSTDSLTPHQPIAQQ-----CSDSLESIPAGQAASDDLDRDVPAGVGASP-----EHA 160
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 154 EAFKQVF---GSLSRRTAQVRYLKYQEEGKVSAYVLRLETLRLKAVEKRAIPRIAD 210
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 161 EPEVQVVPFGSGOIIFLPFTCIGYTATNQD-----FIQLSTLIRQAIE-RQLP----- 207
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 211 QVRLEQVMAGATLNQMLWCLRELKDKQGGPPSFLELMKVIREEEERASFEENESIE---- 266
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 208 -----AWIEAANQREEGQGEEREE-----EDEEEEDVAENRYFEMGPP 248
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 267 EPEERDGYGRWNHGGDD 283
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 249 DVEEEGGGQGEEREE 265
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

```

```

RESULT 8
US-08-922-635-5
; Sequence 5, Application US/08922635A
; Patent No. 6033871
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
; TITLE OF INVENTION: AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6033871
; CURRENT APPLICATION NUMBER: US/08/922,635A
; CURRENT FILING DATE: 1997-09-03
; EARLIER APPLICATION NUMBER: 08/650,766
; EARLIER FILING DATE: 1996-05-20
; EARLIER APPLICATION NUMBER: 60/012,600
; EARLIER FILING DATE: 1996-03-01
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5

```



Best Local Similarity 22.0%; Pred. No. 1.8;  
Matches 82; Conservative 40; Mismatches 121; Indels 130; Gaps 16;

QY 14 TPNDTEFLRLNLFLEKEG-----QTSGMFRALQGVSPATVPCISPELLAHLL 65  
DB 410 SPAQEEEMEEEBE-----EKEGEAGEAESEK-----GEELLPPPESTP1--PANLSQNL 459  
QY 66 QGAMAHAPQLLPMRYKLRVFS----- 88  
DB 460 EAAATQVAVSPKRRKIKELNKEAVGDLDDAFKEANPAVPEVNOPPAGSNPGPSE 519  
QY 89 GSAVPAPEEESFEVWLEQATEI-----VKEW-PVTEAEKKRWLAESLRG 131  
DB 520 GSGVPPRPEADETWDSKEDIHNAENIQGEQKVEYKSDQWKPNNLEKKRYDREFLG 579  
QY 132 -----PALDLMIHIVQ-----DNPSISVECELEAPKQVFGSL----- 163  
DB 580 QFTIPASMQKEGPEGLPHFSDVVDLKANKTLPRLDPTLRQINGCGPDTFPPFANLGRITLS 639  
QY 164 -----ESRTAQV-----RYLKYQEBEGKVSAYVLRLETLRLKAVEKRAIP-- 205  
DB 640 TRGPPRGPGGELPRGQAGLGRPSQGPKEPKIATVMTEDIKLNKAEKANKPSS 699  
QY 206 RRIADQVRLQVWAGATLNMQLMCRRLKDXQGPSPFLMLKVIRE-----ESEE----- 256  
DB 700 KRTAADKDRGEEDADGSK-TQDLPRVRSLNKLTPQMFQQLMKQVTLADTTEERLKV 758  
QY 257 -EASFENESTEEP 268  
DB 759 IDLIFE-KAISEP 770

RESULT 12  
US-08-556-419-25  
; Sequence 25, Application US/08556419C  
; Patent No. 6093549  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Christopher  
; APPLICANT: Li, Xiao-Jiang  
; APPLICANT: Li, Shi-Hua  
; APPLICANT: Sharp, Alan  
; APPLICANT: Lananan, Anthony  
; APPLICANT: Worley, Paul  
; APPLICANT: Snyder, Solomon  
; TITLE OF INVENTION: Huntingtin-associated protein  
; FILE REFERENCE: 01107.52271  
; CURRENT APPLICATION NUMBER: US/08/556,419C  
; CURRENT FILING DATE: 1995-11-09  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Fast-SEQ for Windows Version 3.0  
; SEQ ID NO 25  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-556-419-25

Query Match 6.4%; Score 93.5; DB 3; Length 331;  
Best Local Similarity 22.6%; Pred. No. 0.2;  
Matches 53; Conservative 40; Mismatches 85; Indels 57; Gaps 11;

QY 54 PCTISPELLAH--LLQAMAHAPQLLPMRYKLRVFGSSAVPAPEESFEVWLEQATEIV 111  
DB 44 PCDAPKLISQEAULLHQ--HHCFO--LEALQEKRLLL-----EENHQL-REASQ-- 88  
QY 112 KENPVTEAEKKRWLAESLRGALDMLHIVQADNPSISVECELEAFKQVFGSLRRRTAQV 171  
DB 89 -----LDLTL-----DEQMLILECVQFSEASQMAELSEVLV 122  
QY 172 RYLKTY---QEEGKVSAYVLRLETLRL-----KAVEKRAIPRRIADQVRLQVWAG-- 220  
DB 123 LRLENVERQOQVARIQAQVKLQOORCMYGAETKQLQKLAKEKIQMLQEEETLPGF 182  
QY 221 -ATINQMLWCRRLKDXQGPSPFLMLKVIREEESEASFENESIEPEERDGY 274

DB 183 QETLAELRTSLRRMISD--PVYFMERNYEMPRGDTSSLRDYDFRYSEDRQVRGP 235  
RESULT 13  
US-07-853-913-2  
; Sequence 2, Application US/07853913  
; Patent No. 5338839  
; GENERAL INFORMATION:  
; APPLICANT: McKay, Ronald D.G.  
; APPLICANT: Lendahl, Urban  
; TITLE OF INVENTION: Nestin Expression As An Indicator of  
; TITLE OF INVENTION: Neuroepithelial Tumors  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/853,913  
; FILING DATE: 19920319  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/660,412  
; FILING DATE: 22-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/603,803  
; FILING DATE: 25-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/201,762  
; FILING DATE: 02-JUN-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/180,548  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: MIT-4641AAAA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1805 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-853-913-2

Query Match 6.4%; Score 93; DB 1; Length 1805;  
Best Local Similarity 22.8%; Pred. No. 2.9;  
Matches 68; Conservative 35; Mismatches 137; Indels 58; Gaps 10;

QY 15 PNODTEFLRLNLFLEKEGQTVSGMFRALQGVSPATVPCISPELLAHLLQAMAHAPQ 74  
DB 664 PGADQMLEKL-----VSKEDQSPRSEEDQEAACRFLQKENQSPGLGYEEAGQILE---- 716  
QY 75 PLLPMRYKLRVFGSSAVPAPEESFEVWLEQATEIVKWPVTEAEK-----KRWLA 127  
DB 717 -----RLIEKESQSLRSPPEEDQAGSLQKQEPGLGYEADQMLERLIEKESQ 769  
QY 128 SLRGPDALDMLHIVQADNPSISVECELEAFKQVFGSLRRRTAQVRYLKYVQEEG----- 181  
DB 770 SLKSPENORIGKPLERENQKSLRYLEENQETVPVPLESRNQRPLRSLVEEESQRIKPL 829  
QY 182 EKVSAYVLRLETLRLKAVEKRAIPRIADQVRLQVWAGATLNMQLWCRRLKDXQ-- 238

DB 830 EKV5-----QDS1GSLAENVOPIRYLEE---DDCINKSLTEDKTKHKSLSLEDNGDS 880

QY 239 -----pppfLElMKVIRREEEEAASPENESIEPE--ERDGYRWNHE 280  
||| |::| | | | | | |  
DH 881 TITPORSETOVIAPPER--EDORIVNHLEKESOFRSSREEEOVWERSLEGE-NHE 935

881 TTTPOESETOVST.RPPEE--EDORTVNHLEKESOEFSRSSEEEEOVMERSLEGE-NHE 935

## RESULT 14

```

RES001.14
US-09-252-991A-19048
; Sequence 19048, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19048
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19048

```

Query Match 6.3%; Score 92; DB 4; Length 736;  
Best Local Similarity 22.9%; Pred. No. 0.95;  
Matches 56: Conservative 33; Mismatches 98; Indels 58; Gaps 11;

[illegible]

68 GHWRI,VEAPPSEEEK-----LEGYGLVPFGLFRLLPDGSOPPAPSAASEAETPPPA 119

QY 56 ISPEIIAHLICQAMAHAPQILPMRYKLRVF--SGSAVPAPEESFEVWLEQATEIVKE 113

120 PAPSPTAETAROMGAEA---I.PEKTAGAEFFEGEGSRCSNDODSALAFIRO---VRD 171

QY 114 WPVTEAKKR-----WLAESLRGPAIDLMHIVQADNPISVEECLEAFKQVF 160

Db 172 AC LGEAETKALANSRDLILGACGWEQEELGG-----VLAQGVESAAGKAFATYLEAA 223

[illegible]

224 ANFYSGRFDEAEQGF KALQDVSPWTKETALILQARI DLNAAHQNAFDDMGFEELQNVDK 209

Qy 212 VRLEQ 216



## RESULT 15

US-08-056-200-94  
Sequence 94, Application US/08056200  
Patent No. 5616500  
GENERAL INFORMATION:  
APPLICANT: Steinert, Peter M.  
APPLICANT: Lee, Seung-Chul  
APPLICANT: Kim, In-Gyu  
APPLICANT: Chung, Soo-Il  
APPLICANT: Park, Sang-Chul  
TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and  
TITLE OF INVENTION: Methods of Using Same  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 ; Search time 52.2515 Seconds  
(without alignments)  
1530.308 Million cell updates/sec

Title: US-10-037-860-11

Perfect score: 1462

Sequence: 1 VQKGGVWKVFKTPNQDTE.....SIEEPEERDGYGRWNHEGDD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1459	99.8	283	3	AAB12528 Human Ma4
2	1459	99.8	364	7	ADC08977 Oncogene
3	755	51.6	149	3	AAB12527 Human Ma3
4	628	43.0	353	4	AAB74701 Human mem
5	618.5	42.3	455	5	ABB05727 Human sig
6	618.5	42.3	463	3	AAB12529 Human Ma5
7	618.5	42.3	463	3	AAB42315 Human ORF
8	597	40.8	452	6	AAO16179 Human pro
9	596.5	40.8	351	4	AAE01340 Human gen
10	596.5	40.8	351	4	AAE01340 Human gen
11	596.5	40.8	351	4	AAE01340 Human gen
12	593	40.6	195	3	AAU08664 Human mem
13	552	37.8	329	3	AAB12526 Human Ma2
14	550	37.6	312	3	AAB43023 Human ORF
15	503.5	34.4	280	4	AAE01336 Human gen
16	480.5	32.9	237	4	ADAB94854 Human pro
17	475.5	32.5	399	6	ADAB94854 Human pro
18	475.5	32.5	399	6	ADAB94854 Human pro
19	394	26.9	403	5	ABG99947 Human nov
20	394	26.9	403	5	ABG99947 Human nov
21	394	26.9	403	5	ABG99947 Human nov
22	393	26.9	403	6	ABO14772 Novel hum
23	380.5	26.0	402	4	ABO14772 Novel hum
24	378.5	25.9	337	4	AAB60478 Human cel
25	206.5	14.1	110	4	AAO01787 Human pol

## ALIGNMENTS

## RESULT 1

AAB12528  
ID AAB12528 standard; protein; 283 AA.

XX AAB12528;

DT 02-NOV-2000 (first entry)

DE Human Ma4 protein SEQ ID NO:11.

KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;  
paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;  
breast cancer; parotid gland cancer; lung cancer; testicular cancer;  
germ-cell tumour.

OS Homo sapiens.

XX JP2000146982-A.

PD 26-MAY-2000.

PF 10-NOV-1999; 99JP-00320171.

XX 10-NOV-1998; 98US-00189527.

XX (SLOK ) SLOAN KETTERING INST CANCER RES.

WPI; 2000-468119/41.

XX N-PSDB; AAA60836.

PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic  
encephalitis or neoplasm e.g. colon cancer comprising assessing a test  
sample for the presence or absence of antibodies to a Ma family  
polypeptide.

XX Claim 48; Fig 7-8; 27pp; Japanese.

CC The present invention describes a method for diagnosing a paraneoplastic  
syndrome or neoplasm. The method comprises assessing a test sample for  
the presence or absence of antibodies to a Ma family polypeptide (I). The  
method is used to diagnose a paraneoplastic syndrome especially  
paraneoplastic limbic encephalitis and/or brainstem encephalitis or  
neoplasm in an individual. The method diagnoses the neoplasm by assessing  
antibodies to (i) preferably Ma1, which is indicative presence of breast  
cancer, colon cancer, parotid gland cancer, lung cancer, testicular  
cancer, germ-cell tumours or Ma2, which is indicative of testicular  
cancer, germ-cell tumour, and lung cancer. The present sequence is the  
Ma4 protein as given in the present invention

26 193.5 13.2 615 4 ABG19651 Novel hum  
27 193.5 13.2 615 4 ABG14259 Novel hum  
28 177 12.1 80 4 ABB15036 Human ner  
29 140 9.6 439 6 ABP75736 Human sec  
30 117 8.0 538 4 ABUS2641 Human bra  
31 117 8.0 538 7 ADC31124 Human nov  
32 113 7.7 2383 5 ABG65631 Human bre  
33 113 7.7 2442 3 AAY77575 Human cyt  
34 110.5 7.6 584 4 AAU28186 Novel hum  
35 110.5 7.6 1214 3 AAY57444 Mouse  
36 110.5 7.6 1715 3 AAY57449 Mouse  
37 109 7.5 1270 6 ABO14659 Novel hum  
38 107 7.3 471 4 AAM93676 Human pol  
39 107 7.3 525 5 ABB57353 Mouse isc  
40 107 7.3 543 4 AAM93466 Human pol  
41 107 7.3 543 7 ADE57654 Human pro  
42 107 7.3 595 4 AAB94691 Human pro  
43 107 7.3 648 2 AAY32157 Human SH3  
44 107 7.3 1031 4 AAU87168 Novel cen  
45 107 7.3 1035 4 AAM43519 Human pol

```
XX SQ Sequence 283 AA;
Query Match 99.8%; Score 1459; DB 3; Length 283;
Best Local Similarity 99.6%; Pred. No. 3.9e-132;
Matches 282; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VQKGGVWKVIFKTPNQDTEFLERLNLFLEKQGVSGMFRALQGVSPATVPCISPEL 60
DB 1 VQKGGVWKVIFKTPNQDTEFLERLNLFLEKQGVSGMFRALQGVSPATVPCISPEL 60
QY 61 LAHLGQAMAHAPQPLPMRYKLVFSGSAVPAPEESFEVWLQATEIVKEWPVTEAE 120
DB 61 LAHLGQAMAHAPQPLPMRYKLVFSGSAVPAPEESFEVWLQATEIVKEWPVTEAE 120
QY 121 KKRWLAEISLGRGALDLMHIVQADNPSISVEECLEAFKQVFGSLESRRRTAQVRYLKYQEE 180
DB 121 KKRWLAEISLGRGALDLMHIVQADNPSISVEECLEAFKQVFGSLESRRRTAQVRYLKYQEE 180
QY 181 GEKVSAYVLRLETLRKAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKQGGPP 240
DB 181 GEKVSAYVLRLETLRKAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKQGGPP 240
QY 241 PSFLELMKVIREEEBEASFENESIEPEERDGYGRWNHEGDD 283
DB 241 PSFLELMKVIREEEBEASFENESIEPEERDGYGRWNHEGDD 283
RESULT 2
ADC08977
ID ADC08977 standard; protein; 364 AA.
XX AC ADC08977;
XX DT 18-DEC-2003 (first entry)
XX DE Onconeural antigen Ma2 protein.
XX KW Human; Ma2; onconeural; antigen; Alzheimer's disease;
XX KW neurodegenerative disease; diagnosis; neuroprotective; gene therapy.
XX OS Homo sapiens.
XX PN WO2003073104-A2.
XX XX 04-SEP-2003.
XX XX 26-FEB-2003; 2003WO-EP001946.
XX XX 26-FEB-2002; 2002EP-00004177.
XX XX 26-FEB-2002; 2002US-0359307P.
XX XX (EVOT-) EVOTEC NEUROSCIENCES GMBH.
XX XX Hipfel R, Von Der Kammer H, Pohlner J;
XX XX WPI; 2003-721818/68.
XX XX GENBANK; O94959, KIAA0883.
XX XX
XX PT Diagnosing or prognosticating, or determining increased risk of
XX PT developing a neurodegenerative disease by determining level or activity
XX PT of a transcription or translation product of a gene coding for Ma
XX PT onconeural antigen.
XX XX
XX XX Disclosure; Fig 9; 51pp; English.
XX XX
XX XX The present sequence is the protein sequence of human onconeural
XX XX antigen Ma2. The invention discloses the detection and differential
XX XX expression and regulation of the Ma2 gene in specific brain regions of AD
XX XX patients. The Ma2 gene and its transcription and/or translation products
XX XX may have a causative role in the regional selective neuronal degeneration
XX XX typically observed in AD, or may confer a neuroprotective function to the
XX XX remaining nerve cells. Methods are claimed for diagnosing or
```

```
CC prognosticating a neurodegenerative disease, for monitoring the
CC progression of a neurodegenerative disease, and for evaluating treatment
CC of a neurodegenerative disease, especially AD, in a subject by
CC determining the level and/or activity of a transcription or translation
CC product of an Ma onconeural antigen gene, especially Ma2. Also claimed
CC are: a method for treating or preventing AD and related neurodegenerative
CC disorders using the Ma2 gene or its transcription or translation product;
CC a method of screening for modulating agents of neurodegenerative diseases
CC ; and a recombinant non-human animal comprising an Ma2 gene sequence,
CC which is useful for screening, testing and validating candidate
CC diagnostic and therapeutic agents.
XX SQ Sequence 364 AA;
Query Match 99.8%; Score 1459; DB 7; Length 364;
Best Local Similarity 99.6%; Pred. No. 5.5e-132;
Matches 282; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VQKGGVWKVIFKTPNQDTEFLERLNLFLEKQGVSGMFRALQGVSPATVPCISPEL 60
DB 82 VQKGGVWKVIFKTPNQDTEFLERLNLFLEKQGVSGMFRALQGVSPATVPCISPEL 141
QY 61 LAHLGQAMAHAPQPLPMRYKLVFSGSAVPAPEESFEVWLQATEIVKEWPVTEAE 120
DB 142 LAHLGQAMAHAPQPLPMRYKLVFSGSAVPAPEESFEVWLQATEIVKEWPVTEAE 201
QY 121 KKRWLAEISLGRGALDLMHIVQADNPSISVEECLEAFKQVFGSLESRRRTAQVRYLKYQEE 180
DB 202 KKRWLAEISLGRGALDLMHIVQADNPSISVEECLEAFKQVFGSLESRRRTAQVRYLKYQEE 261
QY 181 GEKVSAYVLRLETLRKAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKQGGPP 240
DB 262 GEKVSAYVLRLETLRKAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKQGGPP 321
QY 241 PSFLELMKVIREEEBEASFENESIEPEERDGYGRWNHEGDD 283
DB 322 PSFLELMKVIREEEBEASFENESIEPEERDGYGRWNHEGDD 364
RESULT 3
AAB12527
ID AAB12527 standard; protein; 149 AA.
XX AC AAB12527;
XX DT 02-NOV-2000 (first entry)
XX DE Human Ma3 protein SEQ ID NO:9.
XX XX
XX KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
XX KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
XX KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
XX KW germ-cell tumour.
XX XX
XX OS Homo sapiens.
XX XX
XX PN JP2000146982-A.
XX XX 26-MAY-2000.
XX XX
XX PF 10-NOV-1999; 99JP-00320171.
XX XX 10-NOV-1998; 98US-00189527.
XX XX
XX XX (SLOK ) SLOAN KETTERING INST CANCER RES.
XX XX
XX XX WPI; 2000-468119/41.
XX XX N-PSDB; AAA60835.
XX XX
XX XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
XX XX encephalitis or neoplasm e.g. colon cancer comprising assessing a test
XX XX sample for the presence or absence of antibodies to a Ma family
XX XX polypeptide.
```





KW	antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW	cholesterol ester storage; systemic lupus erythematosus; infection;
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW	bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW	thrombosis; contraceptive.
XX	
XX	Homo sapiens.
OS	
XX	WO200058473-A2.
PN	
XX	
XX	
PD	05-OCT-2000.
XX	
XX	31-MAR-2000; 2000WO-US008621.
FF	
XX	
PR	31-MAR-1999; 99US-0127607P.
PR	02-APR-1999; 99US-0127636P.
PR	05-APR-1999; 99US-0127728P.
PR	30-MAR-2000; 2000US-00540763.
XX	
XX	(CURA-) CURAGEN CORP.
PA	
XX	
XX	Shimkets RA, Leach M;
PI	
XX	
DR	WPI; 2000-602362/57.
DR	N-PSDB; AAC76524.
XX	
XX	
PT	Novel nucleic acids and peptides derived from open reading frame X,
PT	useful for treating e.g. cancers, proliferative disorders,
PT	neurodegenerative disorders and cardiovascular disease.
XX	
PS	Claim 11; Page 3345-3347; 5507pp; English.

AACT74446 to RAC77606 encode the proteins given in ABA40237 to ABA843397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticonvulsant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive

Sequence 463 AA;

Query Match	42.3%;	Score 618.5;	DB 3;	Length 463;
Best Local Similarity	50.2%;	Pred. No. 1.1e-50;		
Matches 135;	Conservative 44;	Mismatches 85;	Indels 5;	Gaps 3
QY	1	VQGGVWKVFKPTNQDTEFLRLNLFLKEQQTQSGMRALQGGVSPATVPCISPEL	60	
Db	82	IPGKGGPEVIVKPNRSDGFEFLNRLNRLFLSEERATVSDMNRVLGSDTNCSPARVVTISPEF	141	
QY	61	LAHLGQAMAHAPQLL--PMRYKRLYFGSSAVPAPEESFEVLQFQATVKEKWPVTEA	119	
Db	142	WT--WAQTLGAAVQFLLEQMLYRELRYFSGNTSIFGALAFDAWLHHTTEMLQMKNQVPEG	199	
QY	120	EKKRWLAESLRGPALDLMHIVQADNPISIVTEBCLEAFKQVFGSLSSRRRTAQVRYLYKTYOE	179	
Db	200	EKKRLMECLRGPALQVGLGRASNASITVEECALQQVFGFVESHKTAQVKLCKAYOE	259	

```

Qy 180 EBEKYSNVLRLETLRLKAVKRAIPRIADQVLEQVMAGATLNQMLWCLRELKDQGP 239
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 260 AGEKSYSFVLRLEPLQLQAVNNVSVRRNNQTRKLVLSGATLPDLRLDKLMKQRRK 319
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 240 PFSFLEMKVIREEEEEESAF--ENESIE 266
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 320 PGFGLALVKLIREEEWEATLPGDRESLE 348
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 8	
AAO16179	
ID	AAO16179 standard; protein; 452 AA.
XX	
XX	
XX	AAO16179;
XX	
XX	
XX	28-MAR-2003 (first entry)
XX	
XX	Human protein #5.
XX	
XX	
XX	Human; vaccine; adult whole brain; foetal whole brain; tonsil;
KW	adult hippocampus; disease-associated SNP analysis; knockout mouse;
KW	disease model mouse; cancer; neurological disorder.
KW	

OS Homo sapiens.  
XX  
XX -  
PN WO200299103-A1.  
XX  
XX  
PD 12-DEC-2002.  
XX  
XX  
PF 27-MAY-2002; 2002WO-JP005134.  
XX  
XX  
PF 04-JUN-2001; 2001JP-00168370.  
PF 16-AUG-2001; 2001JP-00246915.  
PF

PA	(KAZU-) KAZUSA DNA RES INST FOUND.
PA	(PROT-) PROTEIN EXPRESS CO LTD.
XX	
PI	Ohara O, Nagase T, Nakajima D;
XX	
DR	WPI: 2003-140622/13.
DR	N-PSDB; AAL51207.
XX	
PT	DNA preferentially expressed in human adult and fetal brain tissue
PT	for diagnosis, treatment and analysis of cancer and mental disorder
XX	
PS	Claim 1; Page 56-60; 73pp; Japanese.

The invention comprises the amino acid and coding sequences of several human proteins that are preferentially expressed in adult whole brain, foetal whole brain, tonsil and adult hippocampus tissue. The DNA sequences are useful for the analysis of disease-associated single nucleotide polymorphisms and the production of knockout and human model mice. The DNA and protein sequences of the invention are useful in the prevention (vaccine) and treatment of cancer and neurological disorders. The present amino acid sequence represents a human protein. The invention

[illegible]

Db	194	QVSEVEKRRRLLESURGPALSIIMRVLQANNDSITVEQCLDALKQIFGDKEDFRAQSQREL	253
Qy	175	KTYQEEGKVSAYVLRITLLTKRKVAKRAIPRIADQVRLEOVMGATINQMLWCRLREL	234
Db	254	QTSPIKGKSVSTFLURLBPQLQAVHKSPVSRSTDMIRLKHLLARVAMTPALRGKLELL	313
Qy	235	KDQPPPPFLEIMKVIREEE---EASFENESISPEERDGYGR	276
bh	314	DQRCGCPNPFLEIMKLRIDEEEWENTEAVMKNK-----EKPSGRGR	353

RESULT 9  
AAE01340  
ID AAE01340 standard; protein; 351 AA.  
XX  
XX  
AC AAE01340;  
XX  
XX  
17-JUL-2001 (first entry)  
DT  
DT  
XX  
XX  
DE Human gene 22 encoded secreted protein fragment, SEQ ID NO:205.

Human; secreted protein; proliferative disorder; cancer; tumour;  
fetal abnormality; developmental abnormality; haematopoietic disorder;  
immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
inflammation; allergy; neurological disorder; Alzheimer's disease;  
Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
cardiovascular disorder; angiotensin disorder; kidney disorder;  
gastrointestinal disorder; pregnancy-related disorder;  
endocrine disorder; infection; wound healing; vulnuary; cell culture;  
chemotaxis; food additive; gene therapy; binding partner identification

AA  
 OS Homo sapiens.  
 XX  
 PN WO200134769-A2.  
 XX  
 PD 17-MAY-2001.

XX	01-NOV-2000; 2000WO-US030040.
PF	
XX	
XX	
PR	05-NOV-1999; 99US-0163580P.
PR	30-JUN-2000; 2000US-0215130P.
XX	
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Komatsculis GA, Wei P, Baker KP, Fiscella M;
XX	
XX	
DR	WPI; 2001-308781/32.

XX New isolated nucleic acid molecule encoding a human secreted protein is  
PT used in preventing, treating or ameliorating a medical condition.  
PT  
XX Disclosure: Page 46: 519pp: English.  
PS

XX  
PS Disclosure: Page 46: 519pp; English:

AAD05121-AA0D05203 represent cDNAs corresponding to 24 human secreted  
 protein genes, and AA0E1232-AA0E1311 represent the proteins they encode.  
 AA0E1312-AA0E1340 represent human secreted protein variants or fragments.  
 The secreted proteins and their genes are useful for preventing, treating  
 or ameliorating medical conditions, e.g., by protein or gene therapy.  
 Pathological conditions can be diagnosed by determining the amount of the  
 new protein in a sample or by determining the presence of mutations in  
 the new genes. Specific uses are described for each of the 24 genes,  
 based on the tissues in which they are most highly expressed, and include  
 developing products for the diagnosis or treatment of proliferative  
 disorders, cancer, tumours, foetal and developmental abnormalities,  
 haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 diseases (e.g., rheumatoid arthritis), inflammation, allergies, and  
 neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 pregnancy-related disorders, endocrine disorders, and infections. The



```

Qy 120 EKKRWLABSLRGPAALDLMHIIVQADNPSIVSECELEAFKQVFGSLESRRTAQVRYLKYOE 179
Db 201 EKRRLESLSRGPAALDVRVLKINPLITVDECLQALEEVFGVTDNPRELQVKYLYTTQK 260

Qy 180 EGEKVSAYVRLLETLLRKAVERKRAIPRIADQVRLEQVMAGA---TLNQLMWCRLRELKD 236
Db 261 DEELUSAVVRLLEPLQLKLVQGAERDAVNOARLDQVITAGAVHKTIRREL-----NLPE 315

Qy 237 QGPPEFLELMKVIRE---EEEBEA 258
Db 316 DGPAFGFLQLLVLIKDYEAEBEEA 340

RESULT 12
AAB12526
ID AAB12526 standard; protein; 195 AA.
XX AAB12526;
XX 02-NOV-2000 (first entry)
XX Human Ma2 protein SEQ ID NO:7.
XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
KW germ-cell tumour.
XX Homo sapiens.
XX OS
XX JP2000146982-A.
XX 26-MAY-2000.
XX 10-NOV-1999; 99UP-00320171.
XX 10-NOV-1998; 98US-00189527.
XX (SLOK ) SLOAN KETTERING INST CANCER RES.
XX WPI: 2000-468119/41.
XX N-PSDB; AAB60834.
XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
PT sample for the presence or absence of antibodies to a Ma family
PT polypeptide.
XX Claim 48; Fig 2; 27pp; Japanese.
XX The present invention describes a method for diagnosing a paraneoplastic
CC syndrome or neoplasm. The method comprises assessing a test sample for
CC the presence or absence of antibodies to a Ma family polypeptide (I). The
CC method is used to diagnose a paraneoplastic syndrome especially
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
CC antibodies to (i) preferably Ma1, which is indicative presence of breast
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
CC cancer, and germ-cell tumours or Ma2, which is indicative of testicular
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
CC Ma2 protein as given in the present invention
XX
XX Sequence 195 AA;
Query Match 40.6%; Score 593; DB 3; Length 195;
Best Local Similarity 98.3%; Pred. No. 9.2e-49;
Matches 113; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VQGGGKWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEGVSPATVPCISP 60
Db 81 VQGGGKWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEALSPATVPCISP 140

```

```

Qy 61 LAHLIGQAMAHAPQPLLPMRYRKLVRVFGSVAVPAPERSEFEVWLQATEIVKEWP 115
Db 141 LAHLIGQAMAHAPQPLLPMRYRKLVRVFGSVAVPAPERSEFEVWLQATEIVKEWP 195

RESULT 13
AAB12525
ID AAB12525 standard; protein; 329 AA.
XX AAB12525;
XX 02-NOV-2000 (first entry)
XX Human Ma1 protein SEQ ID NO:4.
XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
KW germ-cell tumour.
XX Homo sapiens.
XX OS
XX JP2000146982-A.
XX 26-MAY-2000.
XX 10-NOV-1999; 99JP-00320171.
XX 10-NOV-1998; 98US-00189527.
XX (SLOK ) SLOAN KETTERING INST CANCER RES.
XX WPI: 2000-468119/41.
XX N-PSDB; AAB60833.
XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
PT sample for the presence or absence of antibodies to a Ma family
PT polypeptide.
XX Claim 48; Fig 1; 27pp; Japanese.
XX The present invention describes a method for diagnosing a paraneoplastic
CC syndrome or neoplasm. The method comprises assessing a test sample for
CC the presence or absence of antibodies to a Ma family polypeptide (I). The
CC method is used to diagnose a paraneoplastic syndrome especially
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
CC antibodies to (i) preferably Ma1, which is indicative presence of breast
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
CC cancer, and germ-cell tumours or Ma2, which is indicative of testicular
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
CC Ma1 protein as given in the present invention
XX
XX Sequence 329 AA;
Query Match 37.8%; Score 552; DB 3; Length 329;
Best Local Similarity 46.8%; Pred. No. 1.8e-44;
Matches 117; Conservative 45; Mismatches 74; Indels 14; Gaps 5;

Qy 3 GKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEGVSPATVPCISP 62
Db 83 GKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEALSPATVPCISP 138

Qy 63 HLIGQAMAHAPQPLL-PMRYRKLVRVFGSVAVPAPERSEFEVWLQATEIVKEWPTEAEK 121
Db 139 EMLNYIDNVIQPLVSIWYKRLTFLSGKHPRAWRGNFDPLHTNEVLEEQVSDVEK 198

Qy 122 KRWLAESLRGPAALDLMHIIVQADNPSIVSECELEAFKQVFGSLESRRTAQVRYLKYOE 181
Db 199 TRRLMESLRGPAADVIRILKSNPNPAITTAELCALHEHVFSGVSSRDAQIKFLNTYQNG 258

Qy 182 EKVSAYVRLLETLLRKAVERKRAIPRIADQVRLEQVMAGA---TLNQLMWCRLRELKD 236

```

DB 259 EKLSAYVIRLEPLQKVEKGAIDKDNVQARLEQVIAGANHSIAIRQLWLTGARE--- 315  
QY 237 QGPPPSFLEL 246  
DB 316 -GGPKPLSV 324  
RESULT 14  
AAB43023  
ID AAB43023 standard; protein; 312 AA.  
XX AAB43023;  
AC AAB43023;  
DT 08-FEB-2001 (first entry)  
DE Human ORFX ORF2787 polypeptide sequence SEQ ID NO:5574.  
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
KW vulnary; antiprosiatic; antiparkinsonian; nootropic; neuroprotective;  
KW immunostimulant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW hypotensive; dermatological; coagulant; vasotropic; antidiabetic;  
KW antiviral; antibacterial; immunosuppressive; antiinflammatory;  
KW antianemic; gene therapy; cancer; proliferative disorder; hypertensive;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive.  
XX Homo sapiens.  
XX WO200058473-A2.  
XX 05-OCT-2000.  
XX 31-MAR-2000; 2000WO-US008621.  
XX 31-MAR-1999; 99US-0127607P.  
XX 02-APR-1999; 99US-0127636P.  
XX 05-APR-1999; 99US-0127728P.  
XX 30-MAR-2000; 2000US-00540763.  
XX (CURA-) CURAGEN CORP.  
XX Shimkets RA, Leach M;  
XX WPT; 2000-602362/57.  
XX N-PSDB; AAC77232.  
XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease.  
XX Claim 11; Page 4759; 5507pp; English.  
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX  
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;  
XX antiprosiatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
XX dermatological; immunosuppressive; antiinflammatory; antibacterial;  
XX antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The  
XX sequences can be used for determining the presence of or predisposition  
XX to, or preventing or treating pathological conditions associated with an  
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX  
XX proteins in gene therapy vectors. The proteins and nucleic acids may be  
XX used to treat cancers, proliferative disorders, neurodegenerative  
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,

CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
SQ Sequence 312 AA;  
Query Match 37.6%; Score 550; DB 3; Length 312;  
Best Local Similarity 50.9%; Pred. No. 2.5e-44;  
Matches 113; Conservative 41; Mismatches 66; Indels 2; Gaps 2;  
QY 1 VQKGGVWKVIFKTPNQDTEFLERLNLFLKEKGQTVSGMFRALGOEGVSPATVPCISPEL 60  
DB 82 IPGKGGIWRVFKPPDPDNTFLSLRNLFLAGMTVGLSRLALGHENGLDPEQGMIPDM 141  
QY 61 LAHLGQAMAHAPQPLLP-MRYKLRVFGSAPPAPEEESPEVWLEQATEIVKEWPVTEA 119  
DB 142 WAPMLAQAL-EALQALQCLYKYLKRVSGRESPEEGEEFGRMWPHFTQMIKAWQVDPV 200  
QY 120 EKXRLAELSLRGPALDLMHIVQADNPISVBECEAFKQVFGSLERRTAQRVYLKTYQE 179  
DB 201 EKRRRLLESRLGPDLDVIRVLKINNPLITVDECLQALEEVFGVDNPRELQVKLYTTQK 260  
QY 180 EGEKUSAVVIRLETLRKAVKRAIPRIADQVRLEQVMAGA 221  
DB 261 DEEXLSAYVIRLEPLQKLVQGAIERDAVNAQARLDQVIAGA 302  
RESULT 15  
AAE01336  
ID AAE01336 standard; protein; 280 AA.  
XX AAE01336;  
XX 17-JUL-2001 (first entry)  
XX Human gene 22 encoded secreted protein fragment, SEQ ID NO:201.  
XX Human; secreted protein; proliferative disorder; cancer; tumour;  
KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
KW skin disorder; psoriasis; sepsis; diabetes; kidney disorder;  
KW cardiovascular disorder; angiogenic disorder; kidney disorder;  
KW gastrointestinal disorder; pregnancy-related disorder;  
KW endocrine disorder; infection; wound healing; vulnary; cell culture;  
KW chemotaxis; food additive; gene therapy; binding partner identification.  
XX Homo sapiens.  
XX WO200134769-A2.  
XX 17-MAY-2001.  
XX 01-NOV-2000; 2000WO-US030040.  
XX 05-NOV-1999; 99US-0163580P.  
XX 30-JUN-2000; 2000US-0215130P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;  
XX WPT; 2001-308781/32.  
XX New isolated nucleic acid molecule encoding a human secreted protein is  
XX used in preventing, treating or ameliorating a medical condition.  
XX Disclosure; Page 46; 519pp; English.  
XX



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	102	10.2		608	2	T03476	conserved hypotheat
2	95	8.5		341	2	C72403	hypothetical prote
3	87.5	8.8		1684	2	T02632	hypothetical prote
4	85	8.5		751	2	F83080	hypothetical prote
5	84.5	8.5		3587	2	I40486	surfactin synthet
6	83.5	8.4		623	2	I64034	hypothetical prote
7	83	8.3		1132	2	T03668	phytochrome B - co
8	83	8.3		1295	2	T17220	hypothetical prote
9	83	8.3		1528	2	T66694	hypothetical prote
10	81.5	8.2		217	2	E69823	conserved hypotheat
11	81.5	8.2		293	2	F69188	conserved hypotheat
12	81.5	8.2		435	1	S18609	ethylene-inducibl
13	81	8.1		330	2	G90256	glutamate dehydrog
14	81	8.1		390	2	A2576	DNA primase, proba
15	81	8.1		457	2	I55976	hypothetical prote
16	81	8.1		502	2	T05246	dihydrolipoamide S
17	81	8.1		555	2	S21766	cytochrome P450 mo
18	81	8.1		609	2	A43906	dihydrolipoamide S
19	80.5	8.1		1076	2	D82083	nuclear phosphopho
20	80	8.0		851	2	AG0482	carbamoyl-phosphat
21	79.5	8.0		461	2	H29350	probable hybrid tw
22	79.5	8.0		859	2	F85680	glycolate oxidase
23	79.5	8.0		859	2	C90853	unknown protein en
24	79.5	8.0		859	2	H90909	probable portal pr
25	79	7.9		316	1	C69855	probable portal pr
26	79	7.9		992	2	T08772	ABC transporter ho
27	78.5	7.9		344	2	C72395	hypothetical prote
28	78.5	7.9		360	2	S49330	transcription regu
29	78.5	7.9		394	2	B72419	serine/threonine k
				394	2		conserved hypotheat



A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA  
 A;Residues: 1-3587 <KUN>  
 A;Cross-references: GB:Z99105; GB:AL009126; NID:G2632457; PIDN:CAB12143.1; PID:G2632635  
 A;Experimental source: strain 168  
 R;Fabret, C.; Quentin, Y.; Guiseppe, A.; Busuttill, J.; Haiech, J.; Denizot, F.  
 Submitted to the EMBL Data Library, March 1993  
 A;Reference number: S46967  
 A;Accession: S46968  
 A;Molecule type: DNA  
 A;Residues: 1-32, 'F', 34-41, 'G', 43-109, 'D', 111-114, 'G', 116-138, 'V', 140-258, 'W', 260-308, 'A', 1756-1914, 'PK', 1917-2138, 'SRL', 2142, 'DSL', 2146-2444, 'Q', 2446-2712, 'H', 2714-2722, 'H', 272  
 A;Cross-references: EMBL:X72672; NID:G516358; PIDN:CAA51233.1; PID:G516360  
 R;Fuma, S.; Fujishima, Y.; Corbelli, N.; D'Souza, C.; Nakano, M.M.; Zuber, P.; Yamane, K.  
 Nucleic Acids Res. 21, 93-97, 1993  
 A;Title: Nucleotide sequence of 5' portion of srfA that contains the region required for  
 A;Reference number: S35517; MUID:93181186; PMID:8441623  
 A;Accession: S35518  
 A;Status: significant sequence differences  
 A;Molecule type: DNA  
 A;Cross-references: EMBL:DJ3262; NID:G216345; PID:G216347  
 A;Experimental source: strain 168 trpC2  
 R;Borchert, S.; Patil, S.S.; Marahiel, M.A.  
 FEMS Microbiol. Lett. 92, 175-180, 1992  
 A;Title: Identification of putative multifunctional peptide synthetase genes using high

A;Reference number: S25658  
 A;Accession: S25658  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 168, 'C', 170-171, 281-283, 514-595, 597-647, 'R', 649-679, 'ETL', 683-693, 'DKR', 697,  
 A;Cross-references: EMBL:X65835; NID:G40202; PIDN:CAA46678.1; PID:G40203  
 A;Experimental source: strain ATCC 21332  
 C;Comment: This protein contains several amino acid-activating domains for the synthesis  
 the amino-terminal region of this protein, appear to be required for the development of  
 C;Genetics:  
 A;Gene: srfAB, srfA2  
 C;Superfamily: surfactin synthetase; acetate-CoA ligase homology; acyl carrier protein h  
 F;Keywords: antibiotic biosynthesis; carrier protein; duplication; ligase; phosphopant  
 F;511-951/Domain: acetate-CoA ligase homology <AC11>  
 F;968-1035/Domain: acyl carrier protein homology <ACPI>  
 F;1036-1481/Domain: repeat <RPT1>  
 F;1542-1995/Domain: acetate-CoA ligase homology <ACL2>  
 F;2013-2081/Domain: acyl carrier protein homology <ACP2>  
 F;2082-2529/Domain: repeat <RPT2>  
 F;2591-3024/Domain: acetate-CoA ligase homology <ACL3>  
 F;3041-3108/Domain: acyl carrier protein homology <ACP3>  
 F;999,2045,3073/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 8.5%; Score 84.5; DB 2; Length 3587;  
 Best Local Similarity 18.9%; Pred. No. 2.9e+02;  
 Matches 44; Conservative 44; Mismatches 82; Indels 63; Gaps 8;  
 QY 11 LMSVDEQKSLMTGIPADFEAEIQVLOETLKSGLRYLLGKIFRKQNNANAVILLE 70  
 Db LSKISGQDIVGSVTAGRTNADVQPMGFNTIAL-----RMEAKQQTFAELLELAK 2399  
 QY 71 DTVSAIPSEV-----QKGGVGVKVFETPNODTFLERLNI----- 107  
 Db QNTLSALEHQEYFFEDLVNQLDPRDMSRNFNVWVTENPKELTLQNLSISPYRAH 2459  
 QY 108 -----FLKEG-----QTVSGMFRALGOEALSPATVPCISPELLAHLGQAMA 150  
 Db QGTSKFDLTGLGFTDENGIGLQLEVATDLFAKETAKWS-----EYVLRLL-KAVA 2509  
 QY 151 HAP-QPLLP-----RYRKLRVSGSNVPAPEESFEVWLEQATEVKEWP 195  
 Db DNPQNQLSLLLVETETEQALLLEAWKGKALPVPDTKTVHQLFEETVQRHKORP 2562

RESULT 6  
 I64034  
 Hypothetical protein H11522 - Haemophilus influenzae (strain Rd Kw20)  
 C;Species: Haemophilus influenzae

C;Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999  
 C;Accession: I64034  
 R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.  
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, C.  
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
 Science 269, 496-512, 1995  
 A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
 A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A;Reference number: A64000; MUID:95350630; PMID:7542800  
 A;Accession: I64034  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-623 <ITGR>  
 A;Cross-references: GB:U32828; GB:I42023; NID:G1574362; PIDN:AAAC23178.1; PID:G1574363;  
 A;Reference number: A64000; MUID:95350630; PMID:7542800

Query Match 8.4%; Score 83.5; DB 2; Length 623;  
 Best Local Similarity 21.7%; Pred. No. 40;  
 Matches 33; Conservative 19; Mismatches 51; Indels 49; Gaps 6;

QY 7 DWCRIMSDVDEQKSLMTGTGIPAD-----FEAEIQVLOETLKSGLRYLLG- 52  
 Db 41 DFMNMQAELFSLDDAGIQPKGRNLNQISLAIRKLSEKGVDFSOQ-LKQADGKYVIGR 99  
 QY 53 -----KIFRKQNNANAVILL-----ELLEDDTVSAIPSE-----VQGGGV 87  
 Db 100 CKSVAEALRTIRPTENCQRILVDAYEGSTAGGGEFVADLQDLITPDGGTCFVVPNNGR 159  
 QY 88 WKVIFKTPNQDTEF-----LERLNLFE 110  
 Db 160 WKRLFSSQLQDTDFGVIGGVADDTTNLNAFLD 191

## RESULT 7

phytochrome B - common tobacco

C;Species: Nicotiana tabacum (common tobacco)  
 C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 21-Jan-2000  
 C;Accession: T03668; T03672  
 R;Kern, R.; Gasch, A.; Deak, M.; Kay, S.A.; Chua, N.H.  
 Plant Physiol. 102, 1363-1364, 1993  
 A;Title: PhB of tobacco, a new member of the photoreceptor family.  
 A;Reference number: Z14996; MUID:94105358; PMID:8278560  
 A;Accession: T03668  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-1132 <KER>  
 A;Cross-references: EMBL:L10114; NID:G295345; PIDN:AAA34092.1; PID:G295346  
 R;Lopez-Juez, E.; Nagatani, A.; Tomizawa, K.; Deak, M.; Kern, R.; Kendrick, R.E.; Furuya  
 Plant Cell 4, 241-251, 1992  
 A;Title: The cucumber long hypocotyl mutant lacks a light-stable PHYB-like phytochrome.  
 A;Reference number: Z14997; MUID:92361250; PMID:1498594  
 A;Accession: T03672  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 457-506, 'S', 508-586, 'Q', 587-1132 <LOP>  
 A;Cross-references: EMBL:M65023; NID:G170286; PIDN:AAA34093.1; PID:G170287  
 C;Genetics:  
 A;Gene: phyB  
 C;Superfamily: phytochrome; phytochrome homology  
 C;Keywords: chromoprotein; photoreceptor; phytochromobilin  
 F;80-589/Domain: phytochrome homology <PHY>  
 F;336/Binding site: phytochromobilin (Cys) (covalent) #status predicted

Query Match 8.3%; Score 83; DB 2; Length 1132;  
 Best Local Similarity 21.6%; Pred. No. 92;  
 Matches 43; Conservative 28; Mismatches 72; Indels 56; Gaps 8;

QY 2 LALLEDWCIMSDVDEQKSLMTGTGIPADFEAEIQVLOETLKSGLRYLLGKIFRKQNA 61  
 Db 586 LLILRDSFKDAEASNSKAV----VHAQLGENELQGI--DELSSVAR-----EMVRLIETA 634  
 QY 62 NAVLL-----ELLEDDTVSAIPSEVQGGGVKVFETPNQDTEFLERLNL 108

635 TAPFADVEGRINGNAKVAETDLV--EEAMGKSLVDVHVKESQETAELK-LFNAL 691  
109 LEKEGOTVSGMFALQGEALSPATV----PCISPELLAHLGLG----- 146  
692 RGEEDKNEIKLRTGPEQLKKAFFVWVWACSSKDYTNVGVFCVGDQVTGQKVVMDKF 751  
147 -----QAMAHAPQPLLP 158  
752 IHTQGDYKAIVHSPNPLIP 770

RESULT 8  
T21720  
hypothetical protein F34D10.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T21720  
R:Kershaw, J.  
submitted to the EMBL Data Library, June 1994  
A:Reference number: Z19464  
A:Accession: T21720  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1295 <WIL>  
A:Cross-references: EMBL:Z34799; PIDN:CAA84316.1; GSPDB:GN00021; CESP:F34D10.6  
A:Map position: 3  
A:Introns: 7/3; 46/3; 109/1; 204/2; 227/3; 266/3; 298/2; 336/1; 397/1; 429/3; 512/2; 597/3

Query Match 8.3%; Score 83; DB 2; Length 1295;  
Best Local Similarity 24.9%; Pred. No. 1.1e+02;  
Matches 49; Conservative 26; Mismatches 60; Indels 62; Gaps 11;

QY 35 QEVLTQELKSGRYRL--GKIFRKQENANAVLLELDTDVSAIPS 79  
DB 442 QNAIYSLVDLGARLDADQGITAEYRQNPGLLRPSSAVSSVLMRSMSTDEFFDPG 501  
QY 80 EVQKGGVW-----KVIFKTPNQDTFELRLNLFLEK---EGQTVSGMF 120  
DB 502 E-QSQGGAGMSKQFLSTFVVPNHKIVFK---KNKFFLNLTITPFYKILLEYVTLDFV 557  
QY 121 RALQGEALSPATVPCISPELLAHLGLQAMAHAPQLLPMRYKLRVFGSSAV----PAPE 176  
DB 558 BAAGRDY-----YLAGISSVHI-KIMLNESEFKKL--FSPFFILNPPNAPS 600  
QY 177 EESFEVLEQAT-EIVK 192  
DB 601 EEQVEEVL--ATGEVLK 615

RESULT 9  
T06694  
hypothetical protein T17F15.240 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 22-Oct-1999  
C:Accession: T06694  
R:Quetier, F.; Choisis, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigou  
submitted to the Protein Sequence Database, April 1999  
A:Reference number: Z15793  
A:Accession: T06694  
A:Molecule type: DNA  
A:Residues: 1-1528 <QUR>  
A:Cross-references: EMBL:AL049658; GSPDB:GN00061; ATSP:T17F15.240  
A:Experimental source: cultivar Columbia; BAC clone T17F15  
C:Genetics:  
A:Gene: ATSP-T17F15.240  
A:Map position: 3  
A:Introns: 823/3; 866/3; 930/3; 947/2; 1051/3; 1094/1; 1161/1; 1193/1; 1253/3; 1325/3; 1411/3

Query Match 8.3%; Score 83; DB 2; Length 1528;

Best Local Similarity 27.4%; Pred. No. 1.3e+02;  
Matches 49; Conservative 26; Mismatches 72; Indels 32; Gaps 10;

QY 28 DPEEABIEQVLOET-LKSLGRYRLGKIFRKQENANAVLLELDTD---VSAIPSEVQG 83  
DB 1174 DDEERFQADLKRKQLSL-----DVYGRNRMTSLKTSLEDNDNWVLSVDVTRESQS 1226

QY 84 KGGWVKVIFKTPNQDTFELRLNLFLEKGGQTV--SGMFALQGEALSPATV-----P 134  
DB 1227 SPCV--AIFGTGLQNE--VGEYCNFLNVIQSLWNLGMFRA---EFLRSSTLHHHFGDP 1279

QY 135 CISPELLAHLGLQAMAHAPQLLPMRYKLRVFGSSAVPAPEESF--EVLWQATEIV 191  
DB 1280 CVVCSLYAIFTALSTASSTRKEFPVAPSSLRIALSNLYP---DSSFFQEQAMNDASEVL 1335

RESULT 10  
E69823  
conserved hypothetical protein yhcQ - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
C:Accession: E69823  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao  
A.; Ehrlich, S.D.; Emmerson, P.F.; Entian, K.D.; Errington, J.; Fabbret, C.; Ferrari, E.;  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hulio, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstara, P.; Tognoni, A.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: E69823  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-217 <KUN>  
A:Cross-references: GB:Z99108; GB:AL009126; NID:g2633055; PIDN:CAB12746.1; PID:el182907  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yhcQ

Query Match 8.2%; Score 81.5; DB 2; Length 217;  
Best Local Similarity 25.0%; Pred. No. 16;  
Matches 46; Conservative 33; Mismatches 76; Indels 29; Gaps 8;

QY 16 EOKSLMVTGIP-----ADFEAEABIEQVLOETLKSGLRYRLGKIFRKQENANAVLLEL 68  
DB 7 QQQSQMKGIPKPKHNGHGHEMDFMHEVLSLTFLVLDQFWMLRQFCKDQ-----LLNI 61

QY 69 LEDTDSVAIPSEVQKGGVWVKVFKT---PNQDTFELRLNLFLEKGGQTVSGMFALQ 125  
DB 62 L-DROHQFITSQY-----NITACEFKTSGSEPSQK-----ATYMKEDNOTVYGMQPSQPK 111

QY 126 ---EALSATVPCISPELLAHLGLQAMAHAPQLLPMRYKLRVFGSSAVPAPEESFEV 182  
DB 112 KPVQSMNDIDDSIISQMLCAKAQA-SMLTWASLEMTNPAVRRLVLSAQIQEYVEMATEI 170

QY 183 MLEQ 186  
DB 171 FLYQ 174

RESULT 11  
F69188  
ethylene-inducible protein - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 17-Nov-2000

C:Accession: F69188  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.  
kt, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A>Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: F69188  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-293 <MTH>  
A:Cross-references: GB:AE000846; GB:AE000666; NID:g2621740; PID:g2621740  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH666  
C:Superfamily: hypothetical protein HI1647

Query Match	8.2%;	Score 81.5;	DB 2;	Length 293;
Best Local Similarity	21.7%;	Pred. No. 23;		
Matches	39;	Conservative 30;	Mismatches 50;	Indels 61; Gaps 9;
Qy	59	ENANAVLLELEDTVSAIPSEVGKGVWVKFTNPQDTEFL-----RLNLF	108	
Dd	34	EDSGAVVALE-----KVPADIRASGGVARM--ADPNKVQEIMDAVSIPVMKVRIGHF	86	
		: : : : :   : : : : :   : : : : :   : : : : :		
Qy	109	LEXEGQTSGMFRALG-----QBALSPA-----TVPICISPELLAHLIQMAH	151	
Dd	87	VEAQ-----VLEALGVDMIDEVLTPADERFHIDKKKFTVPVPCG---ARNLGEALRR	137	
		: : : : :   : : : : :   : : : : :   : : : : :		
Qy	152	APOPLPMPR-----YRKLRVFSGSAPPAPEESFEVW-----LEQATEIVKE	193	
Dd	138	IDSGAAWIRTKRGPGTGNIWEAVRHMKMMSETIREIQNKKEEIEIWSRIEAPLELVRE	197	
		: : : : :   : : : : :   : : : : :   : : : : :		

RESULT 12  
S18609 glutamate dehydrogenase (NADP) (EC 1.4.1.4) - Halobacterium salinarum  
N:Alternate names: Glutamic dehydrogenase; NADP-specific glutamate dehydrogenase  
C:Species: Halobacterium salinarum  
C:Date: 13-Jan-1995 #sequence\_revision 23-Mar-1995 #text\_change 03-Jun-2002  
C:Accession: S18609  
R:Benachemou, N.; Baldacci, G.  
Mol. Gen. Genet. 230, 345-352, 1991  
A:Title: The gene for a halophilic glutamate dehydrogenase: sequence, transcription anal  
A:Reference number: S18609; MUID:92114863; PMID:1766432  
A:Accession: S18609  
A:Molecule type: DNA  
A:Residues: 1-435 <BEN>  
A:Cross-references: EMBL:X63837; NTD:949045; PIDN:CAA45327.1; PID:949046  
A:Note: the source is designated as Halobacterium salinarum  
C:Superfamily: glutamate dehydrogenase (NAD(P)+)  
C:Keywords: NADP; oxidoreductase  
F:126/Binding site: substrate (Lys) #status predicted

```
Query Match      8.2%; Score 81.5; DB 1; Length 435;
Best Local Similarity 25.6%; Pred. No. 37;
Matches 41; Conservative 19; Mismatches 59; Indels 41; Gaps 7;
```

QY	4	LLEDW-CRIMSVDEKSLM-----VTGIPADFEEAIEQVLQETLKSIGRY-----	48
Ddb	249	LLDKWGATIVALSVDNGAWYEPDGDITASVPSHDEEPAVITYADTVISNEELLTLDVDV	308
QY	49	---RLLGKIFRKQENANAVLLLE-----LEDTDVSAIPSEVGQGKVVWK	89
Ddb	309	LIPAALGNVTK-ENAEATAADLVVEGANPTTTSTADSILADRDAVIPDLANAGVTV	367
QY	90	VIF---KTPNQDTEFLERINLFLEKEGQTVCMFALGOE	126
Ddb	368	SYFEWLQDINRWLSRWNDELEAMQAA---WRVAKDE	404

RESULT 13  
G90256

DNA primase, probable [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C:Accession: G90256  
Jong, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan  
arrett, R.A.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: G90256  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-330 <KUR>  
A:Cross-references: GB:AE006641; NID:gl3814232; PIDN:AAK41310.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: SSO1048

	Query Match	8.1%	Score 81;	DB 2;	Length 330;
	Best Local Similarity	26.1%;	Pred. No. 29;	Mismatches	39; Indels 34; Gaps 6;
	Matches	31; Conservative	15;		
Qy	9	CRIMSVDDEKSLM-----VTGIPADFEAEIQEVLQETLKSGLRYRLIGIKFRQEENA	61	:	:
Dd	188	CALLDSDERKEIAEAYVMGIVPGYPGGSENAP-----GWVGRNNGRI	229	:	:
Qy	62	NAVLLLELLEDTVS---AIPSEVOCKGGWVKVIFKTPN-QDTEFLERLNLFLKEQQTV	116	:	:
Dd	230	NGYTIDEQVTIDVKRLIRIPNSLHGKSLG--IVKRPVNLDDEFNFTLSPPF---TGYTI	283	:	:

RESULT 14  
AE2576  
hypothetical protein alr9504 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120ze  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:/date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:/accession: AE2576  
R:/kaneko: T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:/title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:/reference number: AB1807; MUID:21595285; PMID:11759840

Query Match	8.1%; Score 81; DB 2; Length 390;
Best Local Similarity	22.5%; Pred. NO. 36;
Matches	51; Conservative 36; Mismatches 74; Indels 66; Gaps 12;
OY	15 DEQSLMTVGTGPADPEE-----AEIOE-----VLQETLKSGL 46
Db	158 DDWQGLKVVGGDFGEINASMNADI DELENRIKLRQSQGDNALTGKELCHIAEEFPSLK 217
OY	47 -----RYVLLGKIFRKQENANAVILLELLETDVSAIPSEVQGGVGVKVIKFTPNQDTEF 101
Db	218 DECEIAPEWLKGLARRGRKPKIFVIALSQSDSVKAL--GIEGDGAI-----RQNFKY 267
OY	102 LERLNLFLKEKGQTV-----SCMFALQGEALSPATVPCISPELLAH---LLGOA 148
Db	268 L-RLGKFVAVHAKKLGNDAVMNWLQSKYRAMLDD--EPAQMPSIDSTLLVHTLRLDLSRA 324
OY	149 MA--HAPQDILLPMRYKLRVFTSGSAVPAPBEESFEVWLEQATIVKE 193
Db	325 IAGENPPEPPPEMDLTNLPHSLKPAIALLKQEG---WSD--SKLIKE 366

RESULT 15

155976  
 dihydrolipoamide S-acetyltransferase (EC 2.3.1.12), liver - rat (fragment)  
 N:Alternate names: primary biliary cirrhosis autoantigen  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 05-May-2000  
 C:Accession: I55976; 171932  
 R:Gershwin, M.E.; Mackay, I.R.; Sturges, A.; Coppel, R.L.  
 J. Immunol. 138, 3525-3531, 1987  
 A:Title: Identification and specificity of a cDNA encoding the 70 kd mitochondrial antigen  
 A:Reference number: I55976; MUID:87196380; PMID:3571977  
 A:Accession: I55976  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-457 <RES>  
 A:Cross-references: GB:D00092; NID:G220817; PID:AAA41813.1; PID:G220818; GB:M16075; NID:  
 A:Experimental source: liver  
 C:Genetic:  
 A:Genome: nuclear  
 C:Complex: component E2 of pyruvate dehydrogenase complex  
 C:Function:  
 A:Description: catalyzes conversion of acetyl-CoA and dihydrolipoamide to S-acetyldihydrolipoamide  
 A:Pathway: pyruvate metabolism  
 C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology  
 C:Keywords: acetyl-CoA; acyltransferase; coenzyme A; lipoamide; mitochondrion  
 F:48-122/Domain: lipoyl/biotin-binding homology <LEF2>  
 F:169-222/Domain: component E3 binding #status predicted <E3B>

Query Match	8.1%;	Score 81;	DB 2;	Length 457;
Best Local Similarity	25.4%;	Pred. No. 44;		
Matches	36;	Conservative	23;	Mismatches 63; Indels 20; Gaps 5;
38	QY	LOETLSKIGRYRLGLKIFKQENANAVLLE---	LLE-DTDSVAIPSEVQSGGVKVIK	93
47	Dbb	MOIVLPALSPPTMTGTVQWEKKVGEKLSGDLAETIDTKATIGPEVQEEGYLAKLVP		106
94	QY	TPNQDTEFLERLNLFLKEKSGTVS--GMFRALGQALSPATVPCISPELLAHLIGQAMAHA		152
107	Dbb	EGTRDVPGLTFLCIIVEKQEDIAAFADYRPTVETSLKPOAPPPVPPEV-----AAVPEPI		160
153	QY	POPLLPMRYRKLVFSGSAVPA		174
161	Dbb	POPLAP-----TPSAAPA		173

Search completed: September 21, 2004, 13:40:39  
Job time : 16.7398 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 / Search time 5.31599 Seconds  
(without alignments)  
1910.028 Million cell updates/sec

Title: US-10-037-860-7

Perfect score: 996

Sequence: 1 PIALLEDWCRMSVDQKSL.....EESFEVWLFQATEIVKEWP 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	982	98.6	364	1	PM22_HUMAN	Q9u142 homo sapien
2	959	96.3	364	1	PM22_MACFA	Q9gm33 macaca fasc
3	766	76.9	365	1	PM22_MOUSE	Q8bhk0 mus musculus
4	466	47.0	353	1	PM21_RAT	Q8vzh4 rattus norv
5	466	46.8	353	1	PM21_HUMAN	Q8nd90 mus musculus
6	459	46.1	353	1	PM21_MOUSE	Q8clc8 mus musculus
7	436	43.8	351	1	MOP1_HUMAN	Q96dy2 homo sapien
8	431	43.3	351	1	MOP1_MACFA	Q95ki4 macaca fasc
9	431	43.3	352	1	MOP1_MOUSE	Q9erh6 mus musculus
10	91	9.1	2298	1	CU05_HUMAN	Q9v1r5 homo sapien
11	87.5	8.8	3859	1	TRAP_HUMAN	Q9v4a5 homo sapien
12	84.5	8.5	520	1	CET1_CANAL	Q91803 candida alb
13	84.5	8.5	3587	1	SREP2_BACSU	Q04747 bacillus su
14	83.5	8.4	623	1	VPS_HAEIN	P44242 haemophilus
15	83	8.3	1132	1	PHYB_TOBAC	P29130 nicotiana t
16	82.5	8.3	886	1	PHYB_MOUSE	P29130 nicotiana t
17	82	8.2	452	1	IM44_HUMAN	Q43615 homo sapien
18	81.5	8.2	217	1	YHCO_BACSU	Q54601 bacillus su
19	81.5	8.2	293	1	Y666_METTH	Q26762 methanobact
20	81.5	8.2	435	1	DH4_HALSA	P29051 halobacteri
21	81.5	8.2	820	1	BCD2_MOUSE	Q921c5 mus musculus
22	81.5	8.2	1077	1	CARB_VIBPA	Q87sf3 vibrio para
23	81.5	8.2	1077	1	CARB_VIBVU	Q8dem2 vibrio vuln
24	81	8.1	330	1	PRIS_SULSO	Q97283 sulfolobus
25	81	8.1	502	1	C3A_ARATH	P48421 arabidopsis
26	81	8.1	555	1	ODP2_RAT	P08461 rattus norv
27	80.5	8.1	489	1	RNF9_MOUSE	Q9wh5 mus musculus
28	80.5	8.1	559	1	DNL1_PYRKO	Q9hhc4 pyrococcus
29	80.5	8.1	1076	1	RNF1_VIBCH	Q9kph9 vibrio chol
30	80	8.0	421	1	EXJ1_HUMAN	Q92949 homo sapien
31	80	8.0	1130	1	PHYB_SOLTU	P34094 solanum tub
32	79	7.9	1016	1	EM11_HUMAN	Q9y6c2 homo sapien
33	78.5	7.9	360	1	CDKA_HUMAN	Q15131 homo sapien

Q13043 homo sapien  
P06920 human papil  
P33897 homo sapien  
Q913c3 pseudomonas  
Q827h6 salmonella  
P26311 salmonella  
Q9yqx9 anabaena sp  
P34641 caenorhabdi  
P26542 human papil  
P49035 daucus caro  
P46940 homo sapien  
P21263 rattus norv

## ALIGNMENTS

RESULT 1  
PMA2\_HUMAN  
ID PMA2\_HUMAN STANDARD; PRT; 364 AA.  
AC Q9UL42; Q94959; Q95145; Q9UL43;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Paraneoplastic antigen Ma2 (Onconeural antigen MA2) (Paraneoplastic  
neural antigen MA2) (40 kDa neuronal protein).  
GN PMA2 OR MA2 OR KIAA0883.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RP TISSUE=Cerebellum;  
RX MEDLINE=99270611; PubMed=10362822;  
RA Voltz R., Gultekin S.H., Rosenfeld M.R., Gerstner E., Eichen J.,  
Posner J.B., Dalmay J.,  
"A serologic marker of paraneoplastic limbic and brain-stem  
encephalitis in patients with testicular cancer.";  
New Engl. J. Med. 340:1788-1795(1999).  
[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Brain;  
RX MEDLINE=99156230; PubMed=10048485;  
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
"Prediction of the coding sequences of unidentified human genes. XII.  
The complete sequences of 100 new cDNA clones from brain which code  
for large proteins in vitro.";  
DNA Res. 5:355-364(1998).  
[3]  
IDENTIFICATION, AND SUBCELLULAR LOCATION.  
RX MEDLINE=99158179; PubMed=10050892;  
RA Dalmay J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,  
Balmaceda C., Bachelot T., Gerstner E., Eichen J., Frennier J.,  
Posner J.B., Rosenfeld M.R.;  
"Mal, a novel neuron- and testis-specific protein, is recognized by  
the serum of patients with paraneoplastic neurological disorders.";  
Brain 122:27-39(1999).  
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.  
CC -!- TISSUE SPECIFICITY: Brain specific. In some patients suffering  
from cancers, it is also specifically expressed by the testicular  
tumor cells.  
CC -!- MISCELLANEOUS: Antibodies against PMA2 are present in sera from  
patients suffering of paraneoplastic neurological disorders.  
CC -!- SIMILARITY: Belongs to the PMA family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/

FT	DOMAIN	333	338	POLY-GLU.
SC	SEQUENCE	364 AA:	41350 MW:	0CF72210D7EC1524 CRC64;

Query Match 96.3%; Score 959; DB 1; Length 364;  
 Best Local Similarity 95.9%; Pred. No. 2.5e-74;  
 Matches 186; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

3	LALLUEDWCKIRSVDEQKSRWGLFVDELRVAGQ	121
62	NAVLELLEDDTVSAIPSEVCGKGWVKVFKTINQDTEFLERLNLFLEKEGQTVSGMR	122
63	NAVLELLEDDTVSAIPSEVCGKGWVKVFKTINQDTEFLERLNLFLEKEGQTVSGMR	122
122	ALGOEALSPATVPCTISPELLAHLIGQMAHAPOQLLPMRYKLRVFGSGAVPAPEESFE	181

123 ALGHEGMSFAIVPCISPELLAHULGQAFAPAFQDFFFRANERTSUCGATINLEET  
182 VWLEQATEIVKEWP 195

b		183 VWLEQATEIVKWP	196
	RESULT 3		
	MA2_MOUSE		
D	FMA2_MOUSE	STANDARD;	PRT; 365 AA.
C	Q8BHK0;		
T	15-MAR-2004	(Rel. 43, Created)	
T	15-MAR-2004	(Rel. 43, Last sequence update)	
T	15-MAR-2004	(Rel. 43, Last annotation update)	
E	Paraneoplastic antigen Ma2 homolog.		

Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.  
NCBI TaxID:10090.

SEQUENCE FROM N.A.  
STRAIN=C57BL/6J; TISSUE=Brain cortex;  
MEDLINE=22354683; PubMed=12466851;  
Ozekaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
Nakai I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
Schraml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
Blake J.A., Bradt D., Brusica V., Chothia C., Corbani L.E., Cousins S.,  
Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
Nagashima T., Numata K., Okido T., Pavan W.J., Perrega G., Pesole G.,  
petrowsky N., Pillai R., Pontius J.D., Qi D., Ramachandran S.,  
Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,  
Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
Yuan Z., Zvolan M., Zhu Y., Zimmer A., Carninci P., Hayateu N.,  
Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
Birney E., Hayashiraki Y.;  
"Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."  
Nature 420:563-573(2002).  
-!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).  
-!- SIMILARITY: Belongs to the PNMA family.



RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[4]  
RN SEQUENCE OF 162-353 FROM N.A.  
RP TISSUE=Testis;  
RC Pouska A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;  
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
RL Submitter (JUL-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: nuclear; nucleolar. In tumor cells, it is  
CC cytoplasmic.  
CC -1- TISSUE SPECIFICITY: Testis and brain specific. In some patients  
CC suffering from cancers, it is also specifically expressed by the  
CC paraneoplastic tumor cells.  
CC -1- MISCELLANEOUS: Antibodies against PNMA1 are present in sera from  
CC patients suffering of paraneoplastic neurological disorders.  
CC -1- SIMILARITY: Belongs to the PNMA family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF037364; AAD13810.3; -  
CC EMBL; AF320308; AAN05100.1; -  
CC EMBL; BC039577; AAH39577.1; -  
CC EMBL; AL834327; CAD38985.1; -  
CC GenBank; HGNC:9158; PNMA1.  
CC MIM; 604010; -  
CC GO; GO:0005737; C:cytoplasm; TAS.  
CC GO; GO:0005730; C:nucleolus; TAS.  
CC GO; GO:0007417; P:central nervous system development; TAS.  
CC GO; GO:0007283; P:spermatogenesis; TAS.  
CC Antigen; Tumor antigen; Nuclear protein.  
KW DOMAIN 335 341 POLY-GLU.  
FT CONFLICT 136 136  
SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;  
  
Query Match 46.1%; Score 459; DB 1; Length 353;  
Best Local Similarity 47.9%; Pred. No. 2,3e-32;  
Matches 93; Conservative 38; Mismatches 57; Indels 6; Gaps 4;  
  
QY 2 LALLEDCWCRIMSVDEQKSLMTVGIPADFEAEIOVQLTSLGRLYLLGKIFRQENA 61  
DB 3 MTLLEDWCRGMDVNSQALLVWGPVNCDEAEIETLQAAMPQVS-YRVLGRMFWEENA 61  
  
QY 62 NAVLLELLEDDTVSAIPSEVQGGVWVIFKTPNQDTEFLRLMLFLEKEGQTVSGMFR 121  
DB 62 KAALLELTGAVDYAIPREMPGKGLWVFPKPTSDAEFLERLHLFLAREGWTVDVAR 121  
  
QY 122 ALGQEALESATVPICSPPELLLAHLGQAMAHAPQPLL-PMYRKRLRVFSSAVPAPEESF 180  
DB 122 VLGFQ--NPPTP--GPENPAEMLNVLNDVQLPLVESIWKYKLTFLSGRDIPGGEETF 177  
  
QY 181 EYVLEQATEIVKEW 194  
DB 178 DPWLEHTNEVLEW 191  
  
RESULT 6  
ID PNMA1 MOUSE STANDARD; PRT; 353 AA.  
AC O8C1C8; O9CVP2;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Paraneoplastic antigen Maf homolog.  
GN PNMA1 OR MAF.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;  
RX MEDLINE=22354683; PubMed=12466851;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kiyosawa H.,  
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Shonbach C., Gojohori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Gough J.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Jarvis E.D.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., King B.D.,  
RA Kanai A., Kawai H., Kawasawa Y., Kedzierzki R.M., King B.D.,  
RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Meglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).  
CC -1- SIMILARITY: Belongs to the PNMA family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AK017476; BAB30762.1; -  
CC EMBL; AK028331; BAC25895.1; -  
CC MGD; MGI:2180564; Pnmal.  
CC Nuclear protein.  
KW DOMAIN 336 341 POLY-GLU.  
FT CONFLICT 136 136  
SQ SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;  
  
Query Match 46.1%; Score 459; DB 1; Length 353;  
Best Local Similarity 47.9%; Pred. No. 8,9e-32;  
Matches 93; Conservative 36; Mismatches 59; Indels 6; Gaps 4;  
  
QY 2 LALLEDCWCRIMSVDEQKSLMTVGIPADFEAEIOVQLTSLGRLYLLGKIFRQENA 61  
DB 3 MTLLEDWCRGMDVNSQALLVWGPVNCDEAEIETLQAAMPQVS-YRVLGRMFWEENA 61  
  
QY 62 NAVLLELLEDDTVSAIPSEVQGGVWVIFKTPNQDTEFLRLMLFLEKEGQTVSGMFR 121  
DB 62 KAALLELTGAVDYSLIPREMPGKGLWVFPKPTSDAEFLERLHLFLAREGWTVDVAR 121  
  
QY 122 ALGQEALESATVPICSPPELLLAHLGQAMAHAPQPLL-PMYRKRLRVFSSAVPAPEESF 180  
DB 122 VLGFQ--NPAP--GPETPAEMLNVLNDVQLPLVESIWKYKLTFLSGKDIPGGEETF 177  
  
QY 181 EYVLEQATEIVKEW 194  
DB 178 DSWLEHSENEIIEW 191

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC -!- SIMILARITY: Belongs to the PNMA family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB060854; BAB46873.1; -.  
 DR Apoptosis. 120 127 BH3-LIKE.  
 FT SITE 336 POLY-GLU.  
 FT DOMAIN 351 AA; 39623 MW; C7530E4496A6FFB3 CRC64;  
 SQ SEQUENCE 351 AA; 39623 MW; C7530E4496A6FFB3 CRC64;  
 Query Match 43.3%; Score 431; DB 1; Length 351;  
 Best Local Similarity 46.4%; Pred. No. 2.1e-29;  
 Matches 90; Conservative 30; Mismatches 72; Indels 2; Gaps 2;  
 QY 2 LALLEDCWCRMSVDEQKSLMTGIPADFEAEIOEVLQETLSLGRVRLGKIFRKQENA 61  
 DB 3 LRLLEDWCRCMDNPKRALLIAGISQSCVAEIEALQAGLAPLGEYRLGLGRFRDRDENR 62  
 QY 62 NAVLELLEDTDVSAIPSEVQGGVWKVIFKTPNQDTFELRLNLFLEKEGQTVSGMFR 121  
 DB 63 KVALVGLTATSHALVPKRIIPGKGGIWRVIFKPPDSNTFLSLRNEFLAGEGWTGELTR 122  
 QY 122 ALGOEALSPATVPCISPPELLAHLGOMAHAPQPLP-MRYRKLRFVPSGSAVPAPEBESF 180  
 DB 123 ALAHENSLDLEQGMIFEMWAPMLAQL-BALQALQCLKYKLRVPSGREPPPGREEF 181  
 QY 181 EVMLEQATEIVKEW 194  
 DB 182 GRWVFHTTOMIKAW 195  
 RESULT 9  
 MOPI MOUSE STANDARD; PRT; 352 AA.  
 ID MOPI MOUSE  
 AC Q9BRH6;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Modulator of apoptosis 1 (MAP-1).  
 GN MOAP1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OC NCBI\_TaxID=10090;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21264738; PubMed=11060313;  
 RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevoit M., Ang K.C.,  
 RA Yu V.C.;  
 RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that  
 RT associates with Bax through its Bcl-2 homology domains."  
 RL J. Biol. Chem. 276:2802-2807(2001).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glasziou C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bernaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Willing L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6Ncr; TISSUE=Hematopoietic, and Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,  
 RA Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hong L.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh F.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitig M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Mediates caspase-dependent apoptosis.  
 CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).  
 CC -!- DOMAIN: The BH3-like domain is required for association with BAX  
 CC and for mediating apoptosis. The three BH domains (BH1, BH2, and  
 CC BH3) of BAX are all required for mediating protein-protein  
 CC interaction (By similarity).  
 CC -!- SIMILARITY: Belongs to the PNMA family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF305551; AAG31787.1; -.  
 DR EMBL; AK019599; BAB31810.1; -.  
 DR EMBL; BC014715; AAH14715.1; -.  
 DR EMBL; BC055374; AAH55374.1; -.  
 DR MGI; MGI:191555; Moap1.  
 KW Apoptosis.  
 FT SITE 120 127 BH3-LIKE.  
 FT DOMAIN 335 340 POLY-GLU.  
 FT CONFLICT 57 57 R -> K (IN REF. 3; AAH55374).  
 SQ SEQUENCE 352 AA; 39404 MW; 8F4630D080495D98 CRC64;  
 Query Match 43.3%; Score 431; DB 1; Length 352;  
 Best Local Similarity 44.9%; Pred. No. 2.2e-29;  
 Matches 89; Conservative 35; Mismatches 66; Indels 8; Gaps 3;  
 QY 2 LALLEDCWCRMSVDEQKSLMTGIPADFEAEIOEVLQETLSLGRVRLGKIFRKQENA 61  
 DB 3 LRLLEDWCRCMDNPKRALLIAGISQSCVAEIEALQAGLAPLGEYRLGLGRFRDRDENR 62  
 QY 62 NAVLELLEDTDVSAIPSEVQGGVWKVIFKTPNQDTFELRLNLFLEKEGQTVSGMFR 121  
 DB 63 NVALICLVETGSLVPKIEIPAKGVRVIFKPPDSDFLCRLNEFLKGEQWTGELTR 122

QY 122 ALGQE-----ALSPATVFCISPELLAHLIGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPE 176  
 Db 123 VLGNRPDLGLDPG---IMPEIRAPMLAQALNEALXPTLQYLRKLSVFGSRDPGPG 179  
 QY 177 EESFVWLEQATEIVKSW 194  
 Db 180 EEFESWMTFTSQVMKTM 197

RESULT 10  
 CU05 HUMAN STANDARD; PRT; 2298 AA.  
 AC Q9Y3R5; Q9Y3Z3;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Protein C21orf5.  
 GN C21ORF5 OR KIAA0933.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20408884; PubMed=10950924;  
 RA Guipponi M., Brunschwag K., Chamoun Z., Scott H.S., Shibuya K.,  
 RA Kudoh J., Delezoide A.L., El Samadi S., Chettouh Z., Rossier C.,  
 RA Shimizu N., Mueller F., Delabar J.M., Antonarakis S.E.;  
 RT "C21orf5, a novel human chromosome 21 gene, has a Caenorhabditis  
 RT elegans ortholog (pad-1) required for embryonic patterning.";  
 RL Genomics 68:30-40 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Shibuya K., Kudoh J., Minoshima S., Kawasaki K., Nakatoh E.,  
 RA Shintani A., Asakawa S., Shimizu N.;  
 RT "Genomic sequencing of 1.2-Mb region on human chromosome 21q22.2.";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953;  
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,  
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,  
 RA Soeda E., Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,  
 RA Pooley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
 RA Reichwald A., Rump A., Schillhabel M., Schudy A., Zimmermann W.,  
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,  
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,  
 RA Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P.,  
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,  
 RA Ramser J., Beck A., Klages S., Hennig S., Riessmann L., Dagand E.,  
 RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
 RA Leirach H., Reinhardt R., Yaspo M.-L.;  
 RT "The DNA sequence of human chromosome 21.";  
 RL Nature 405:311-319 (2000).  
 RN [4]  
 RP SEQUENCE OF 152-2298 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22158633; PubMed=12168954;  
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;  
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual  
 RT curation of 330 KIAA cDNA clones.";  
 RL DNA Res. 9:99-106 (2002).  
 RN [5]  
 RP SEQUENCE OF 677-2298 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99246063; PubMed=10231032;  
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XIII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 6:63-70 (1999).  
 CC -!- TISSUE SPECIFICITY: Ubiquitous.

CC -!- SIMILARITY: TO C.ELEGANS Y18D10A.15.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ237839; CAB41415.1; -.  
 DR EMBL; AP000692; BAA89431.1; -.  
 DR EMBL; AP000699; BAA89431.1; JOINED.  
 DR EMBL; AP000890; BAA89431.1; JOINED.  
 DR EMBL; AP000891; BAA89431.1; JOINED.  
 DR EMBL; AP001725; BAA95548.1; -.  
 DR EMBL; AB023150; BAA76777.2; -.  
 DR Genew; HGNC:1291; C21orf5.  
 DR MIM; 604803; -.  
 DR GO; GO:0007275; P.development; NAS.  
 DR InterPro; IPR007249; Dopey\_N.  
 DR Pfam; PF04118; Dopey\_N; 1.  
 DR FT CONFLICT 498 498 Q -> H (IN REF. 1).  
 FT CONFLICT 1118 1118 G -> C (IN REF. 2 AND 3).  
 FT CONFLICT 1149 1149 H -> P (IN REF. 2 AND 3).  
 FT CONFLICT 1318 1318 L -> P (IN REF. 1).  
 FT CONFLICT 1886 1886 V -> A (IN REF. 1).  
 FT CONFLICT 1974 1974 E -> G (IN REF. 1).  
 FT CONFLICT 2139 2139 G -> E (IN REF. 4).  
 FT CONFLICT 2154 2154 K -> R (IN REF. 1).  
 SQ SEQUENCE 2298 AA; 258221 MW; 2C49AF550739F0 CRC64;

Query Match 9.1%; Score 91; DB 1; Length 2298;  
 Best Local Similarity 22.7%; Pred. No. 19;  
 Matches 53; Conservative 33; Mismatches 89; Indels 58; Gaps 11;

QY 1 PLALLE-----DWCEIMSVDEQKSLMTVGIPADFEAEIQEVLTSLKSLGRYRLGKI 54  
 Db 1579 PUTLLEGLTTISHFCLLEQANQKTKMAAGDPANLRNAR-NAILELPRTVTMTALLWNV 1637  
 QY 55 FRKQENANAVILLELELTDVSAIPSEVOGKGVKVIKTPNQD-TEFLERINLFL---- 109  
 Db 1638 LRKEHTQKRP-VDLGAT-----KGSVVYFKTKTIQKILDFLNPJTAHLGVOL 1687  
 QY 110 -----EKEGQTVSGMFRALGQALSPAT-----VPCISPELLAHLIGQAMAHAP 153  
 Db 1688 TAAVAAVMSRKAQRHSM-KIIPITASASQLTLDVLCALSTLQTDTLHLVKEVVKRP 1746  
 QY 154 Q-----PLLPWRYKRLRVFSGSAVPAPE-BESFVWLEQATEIVK 192  
 Db 1747 QVKGDEKSLVDIPVLOQYAFLO-----RLFPALQENFSSLLGVLKESVQ 1794

RESULT 11  
 TRAP HUMAN  
 ID TRAP HUMAN STANDARD; PRT; 3859 AA.  
 AC Q9Y4A5; Q9Y5218; Q9Y631; Q9Y6H4;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Transformation/transcription domain-associated protein (350/400 kDa  
 DE PCAP-associated factor) (PAF350/400) (STAF40) (Tral homolog).  
 GN TRRAP OR PAF400.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), PARTIAL SEQUENCE, FUNCTION,  
 RP SUBCELLULAR LOCATION, AND INTERACTION WITH MYC AND E2F1.  
 RC TISSUE=Cervical carcinoma;  
 RX MEDLINE=98372443; PubMed=9708738;  
 RA McMahon S.B., Van Buskirk H.A., Dugan K.A., Copeland T.D., Cole M.D.;

RT "The novel ATM-related protein TRRAP is an essential cofactor for the  
RT c-Myc and E2F oncoproteins.";  
RL Cell 94:363-374(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1), SEQUENCE OF 41-90; 337-344; 369-387;  
RP 879-892; 997-1005; 1171-1184; 1237-1247; 1545-1560; 1815-1820;  
RP 1922-1934; 2211-2218; 2260-2275; 2534-2547; 2583-2594; 2706-2726;  
RP 2830-2844; 3567-3573; 3583-3598; 3604-3614; 3712-3730 AND 3822-3834,  
RP AND IDENTIFICATION IN THE PCAF COMPLEX WITH TADA2L; TADA3L; TAF5L;  
RP TAF6L; TAF10; SUPT3H; TAF12 AND TAF9.  
RC TISSUE=Petal heart;  
RX MEDLINE=99102960; PubMed=9885574;  
RX Vassilev A., Yamauchi J., Kotani T., Prives C., Avantaggiati M.L.,  
RA Qin J., Nakatani Y.;  
RA "The 400 kDa subunit of the PCAF histone acetylase complex belongs to  
RT the ATM superfamily.";  
RL Mol. Cell 2:869-875(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22737999; PubMed=12853948;  
RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,  
RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,  
RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,  
RA Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,  
RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,  
RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,  
RA Ozerky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,  
RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,  
RA Kozlowski-Reilly A., Leonard S., Rohlfing T., Rock S.M.,  
RA Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strommatt C.,  
RA Latrelle P., Miller N., Johnson D., Murray J., Woessner J.P.,  
RA Wendt M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,  
RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,  
RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,  
RA Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,  
RA Gillett W., Zhou Y., James R., Phelps K., Iadamoto S., Bubb K.,  
RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Pury T.S.,  
RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,  
RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,  
RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,  
RA Waterston R.H., Wilson R.K.;  
RA "The DNA sequence of human chromosome 7.";  
RL Nature 424:157-164(2003).  
RN [4]  
RP SEQUENCE OF 21-38; 882-893; 2005-2015; 2049-2059; 2063-2074 AND  
RP 3583-3598, AND IDENTIFICATION IN THE STAGA COMPLEX WITH SF3B3;  
RP GCN5L2; KIAA0764; TAF5L; TAF6L; TADA3L; TAF10; TAF12 AND TAF9.  
RX MEDLINE=21448975; PubMed=11564863;  
RA Martinez E., Palhan V.B., Tjernberg A., Lymar E.S., Gamper A.M.,  
RA Kundu T.K., Chait B.T., Roeder R.G.;  
RA "Human STAGA complex is a chromatin-acetylating transcription  
RT coactivator that interacts with pre-mRNA splicing and DNA  
RT damage-binding factors in vivo.";  
RL Mol. Cell. Biol. 21:6782-6795(2001).  
RN [5]  
RP IDENTIFICATION IN THE TFC-HAT COMPLEX WITH TAF5L; TAF6L; TADA3L;  
RP SUPT3H; TAF2; TAF4; TAF5; GCN5L2 AND TAF10.  
RX MEDLINE=99303588; PubMed=10373431;  
RA Brand M., Yamamoto K., Staub A., Tori L.;  
RA "Identification of TATA-binding protein-free TAFII-containing complex  
RT subunits suggests a role in nucleosome acetylation and signal  
RT transduction.";  
RL J. Biol. Chem. 274:18285-18289(1999).  
RN [6]  
RP IDENTIFICATION IN THE TIP60 HAT COMPLEX WITH HTATIP; RUVBL1 AND  
RP RUVBL2.  
RX MEDLINE=20419292; PubMed=10966108;  
RA Ikura T., Ogryzko V.V., Grigoriev M., Groisman R., Wang J.,  
RA Horikoshi M., Scully R., Qin J., Nakatani Y.;  
RT "Involvement of the TIP60 histone acetylase complex in DNA repair and  
RT apoptosis.";  
RL Cell 102:463-473(2000).  
RN [7]

RP INTERACTION WITH GCN5L2.  
RX MEDLINE=20079268; PubMed=10611234;  
RA McMahon S.B., Wood M.A., Cole M.D.;  
RT "The essential cofactor TRRAP recruits the histone acetyltransferase  
RT hGCN5 to c-Myc.";  
RL Mol. Cell. Biol. 20:556-562(2000).  
RN [8]  
RP INTERACTION WITH E2F1 AND E2F4, AND FUNCTION.  
RX MEDLINE=21413878; PubMed=11418595;  
RX Lang S.E., McMahon S.B., Cole M.D., Hearing P.;  
RT "E2F transcriptional activation requires TRRAP and GCN5 cofactors.";  
RL J. Biol. Chem. 276:32627-32634(2001).  
RN [9]  
RP DOMAIN.  
RX MEDLINE=21338486; PubMed=11445536;  
RA Park J., Kunjibettu S., McMahon S.B., Cole M.D.;  
RT "The ATM-related domain of TRRAP is required for histone  
RT acetyltransferase recruitment and Myc-dependent oncogenesis.";  
RL Genes Dev. 15:1619-1624(2001).  
RN [10]  
RP FUNCTION, AND INTERACTION WITH TP53.  
RX MEDLINE=22133580; PubMed=12138177;  
RA Ard P.G., Chatterjee C., Kunjibettu S., Adside L.R., Gralinski L.R.,  
RA McMahon S.B.;  
RT "Transcriptional regulation of the mdm2 oncogene by p53 requires TRRAP  
RT acetyltransferase complexes.";  
RL Mol. Cell. Biol. 22:5650-5661(2002).  
RN [11]  
RP IDENTIFICATION IN THE BAF53 COMPLEX WITH BAF53A; RUVBL1 AND SMARCA4.  
RX MEDLINE=21829261; PubMed=11839798;  
RA Park J., Wood M.A., Cole M.D.;  
RT "BAF53 forms distinct nuclear complexes and functions as a critical  
RT c-Myc-interacting nuclear cofactor for oncogenic transformation.";  
RL Mol. Cell. Biol. 22:1307-1316(2002).  
RN [12]  
RP FUNCTION.  
RX MEDLINE=22628784; PubMed=12743606;  
RA Lang S.E., Hearing P.;  
RT "The adenovirus E1A oncoprotein recruits the cellular TRRAP/GCN5  
RT histone acetyltransferase complex.";  
RL Oncogene 22:2836-2841(2003).  
RN [13]  
RP FUNCTION.  
RX MEDLINE=22651123; PubMed=12660246;  
RA Liu X., Tesfai J., Eward Y.A., Dent S.Y.R., Martinez E.;  
RT "c-Myc transformation domain recruits the human STAGA complex and  
RT requires TRRAP and GCN5 acetylase activity for transcription  
RT activation.";  
RL J. Biol. Chem. 278:20405-20412(2003).  
RN [14]  
RP FUNCTION: Adapter protein, which is found in various multiprotein  
CC chromatin complexes with histone acetyltransferase activity (HAT),  
CC which gives a specific tag for epigenetic transcription  
CC activation. Plays a central role in MYC (c-Myc) transcription  
CC activation, and also participates in cell transformation by MYC.  
CC Required for TP53/p53-, E2F1- and E2F4-mediated transcription  
CC activation. Also involved in transcription activation mediated by  
CC the adenovirus E1A, a viral oncoprotein that deregulates  
CC transcription of key genes. Probably acts by linking transcription  
CC factors such as E1A, MYC or E2F1 to HAT complexes such as STAGA  
CC thereby allowing transcription activation. Probably not required  
CC in the steps following histone acetylation in processes of  
CC transcription activation. May be required for the mitotic  
CC checkpoint and normal cell cycle progression.  
CC SUBUNIT: Interacts with TP53. E2F1 and E2F4 transcription factors.  
CC Interacts directly with MYC. E2F1 and E2F4 transcription factors.  
CC Interacts with GCN5L2. Interacts with GCN5L2. Component of  
CC various HAT complexes. Component of the PCAF complex, at least  
CC composed of TADA2L/ADA2, SUPT3H, TADA3L/ADA3, TAF5L/PAF65beta,  
CC TAF6L/PAF65alpha, TAF10/TAF1130, TAF12/TAF1120, TAF9/TAF1131 and  
CC TRRAP. Component of the TFC-HAT complex, at least composed of  
CC TAF5L, TAF6L, TADA3L, SUPT3H/SPT3, TAF2/TAF1150, TAF4/TAF1135,  
CC TAF5/TAF1100, GCN5L2/GCN5, TAF10 and TRRAP. Component of the  
CC TIP60 HAT complex, at least composed of HTATIP/TIP60,  
CC RUVBL1/TIP49, RUVBL2/TIP48 and TRRAP, which preferentially



RA Fuma S., Fujishima Y., Corbell N., D'Souza C., Nakano M.M.,  
RA Zuber P., Yamane K.;  
RT "Nucleotide sequence of 5' portion of srfa that contains the region  
RT required for competence establishment in *Bacillus subtilis*.";  
RL Nucleic Acids Res. 21:93-97(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / JH642;  
RX MEDLINE=93360813; PubMed=8355609;  
RA Cosmina P., Rodriguez F., de Ferri F., Grandi G., Perego M.,  
RA Venena G., van Sinderen D.;  
RT "Sequence and analysis of the genetic locus responsible for surfactin  
RT synthesis in *Bacillus subtilis*.";  
RL Mol. Microbiol. 8:821-831(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=97124189; PubMed=8969502;  
RA Yamane K., Kumano M., Kurita K.;  
RT "The 25 degrees-36 degrees region of the *Bacillus subtilis*  
RT chromosome: determination of the sequence of a 146 Kb segment and  
RT identification of 113 genes.";  
RL Microbiology 142:3047-3056(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.M., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.J., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Gim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holbappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parzo V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowaka A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
RT *subtilis*.";  
RL Nature 390:249-256(1997).  
RN [5]  
RP SEQUENCE OF 514-800 FROM N.A.  
RC STRAIN=ATCC 21332;  
RX MEDLINE=92290255; PubMed=1601288;  
RA Borchert S., Patel S.S., Marahiel M.A.;  
RT "Identification of putative multifunctional peptide synthetase genes  
RT using highly conserved oligonucleotide sequences derived from known  
RT synthetases.";  
RL FEMS Microbiol. Lett. 71:175-180(1992).  
CC -1- FUNCTION: THIS PROTEIN IS A MULTIFUNCTIONAL ENZYME ABLE TO  
CC ACTIVATE AND POLYMERIZE THE AMINO ACIDS LEU, GLU, ASP AND VAL.  
CC ACTIVATION SITES FOR THESE AA CONSIST OF INDIVIDUAL DOMAINS.  
CC -1- COFACTOR: Contains 3 covalently bound phosphopantetheines.  
CC -1- PATHWAY: Cyclic peptide antibiotic surfactin biosynthesis.  
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme  
CC family.  
CC -1- SIMILARITY: Contains 3 acyl carrier domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: D13262; BAA02523.1; -;  
CC EMBL: X70356; CAA49817.1; -;  
CC EMBL: D50453; BAA08983.1; -;  
CC EMBL: Z99105; CAB12143.1; -;  
CC EMBL: X65835; CAA46678.1; -;  
CC PIR: I40486; I40486.  
CC HSSP: P14687; I40486.  
CC Subtilisin; BG10169; srfAB.  
CC InterPro: IPR000873; AMP-bind.  
CC InterPro: IPR001242; Condensatn.  
CC InterPro: IPR006163; Pp-bind.  
CC InterPro: IPR006162; Ppantne S.  
CC Pfam: PF00501; AMP-binding; 3.  
CC Pfam: PF00668; Condensation; 4.  
CC Pfam: PF00550; pp-binding; 3.  
CC PRINTS: PR00154; AMPBINDING.  
CC PROSITE: PS00012; PHOSPHOPANTHETHEINE; 3.  
CC PROSITE: PS00455; AMP BINDING; 3.  
CC PROSITE: PS00075; ACP\_DOMAIN; 3.  
CC Ligase; Anticibiotic biosynthesis; Phosphopantetheine; Sporulation;  
CC Multifunctional enzyme; Repeat; Complete proteome.  
CC REPEAT ? 1040 DOMAIN 1 (VAL-ACTIVATING).  
CC REPEAT ? 2096 DOMAIN 2 (ASP-ACTIVATING).  
CC REPEAT ? 3114 DOMAIN 3 (D-LEU-ACTIVATING).  
CC REPEAT ? 1036 ACYL CARRIER (ACP) 1.  
CC DOMAIN 2015 2082 ACYL CARRIER (ACP) 2.  
CC DOMAIN 3043 3109 ACYL CARRIER (ACP) 3.  
CC BINDING 999 999 PHOSPHOPANTHETHEINE (POTENTIAL).  
CC BINDING 2045 2045 PHOSPHOPANTHETHEINE (POTENTIAL).  
CC BINDING 3073 3073 PHOSPHOPANTHETHEINE (POTENTIAL).  
CC CONFLICT 33 33 S -> F (IN REF. 1).  
CC CONFLICT 42 42 A -> G (IN REF. 1).  
CC CONFLICT 110 110 Q -> D (IN REF. 1).  
CC CONFLICT 113 115 RQA -> AQG (IN REF. 1).  
CC CONFLICT 139 139 A -> V (IN REF. 1).  
CC CONFLICT 259 259 L -> W (IN REF. 1).  
CC CONFLICT 309 309 R -> A (IN REF. 1).  
CC CONFLICT 478 480 TPA -> SRP (IN REF. 1).  
CC CONFLICT 596 596 MISSING (IN REF. 5).  
CC CONFLICT 648 648 A -> R (IN REF. 1).  
CC CONFLICT 680 682 RHV -> ETL (IN REF. 1).  
CC CONFLICT 694 698 EQSIT -> DKRIS (IN REF. 5).  
CC CONFLICT 788 788 M -> L (IN REF. 5).  
CC CONFLICT 939 940 EL -> LV (IN REF. 1).  
CC CONFLICT 1038 1038 N -> I (IN REF. 1).  
CC CONFLICT 1133 1133 H -> Q (IN REF. 1).  
CC CONFLICT 1310 1310 V -> C (IN REF. 1).  
CC CONFLICT 1333 1333 G -> V (IN REF. 1).  
CC CONFLICT 1384 1384 P -> R (IN REF. 1).  
CC CONFLICT 1582 1582 G -> E (IN REF. 1).  
CC CONFLICT 1677 1682 KRRADG -> E (IN REF. 1).  
CC CONFLICT 1700 1700 C -> S (IN REF. 1).  
CC CONFLICT 1755 1755 F -> K (IN REF. 1).  
CC CONFLICT 1787 1787 T -> S (IN REF. 1).  
CC CONFLICT 1801 1822 GAIAGRVLPDPAFAKPTTG -> AFSPGGLICWRCIC  
CC ETPDNR (IN REF. 1).  
CC LG -> PK (IN REF. 1).  
CC R -> C (IN REF. 1).  
CC A -> V (IN REF. 1).  
CC ARLTP -> LRSLN (IN REF. 1).  
CC E -> Q (IN REF. 1).  
CC ATDLF -> RQICS (IN REF. 1).  
CC CONFLICT 1915 1916  
CC CONFLICT 2075 2075  
CC CONFLICT 2079 2079  
CC CONFLICT 2141 2145  
CC CONFLICT 2445 2445  
CC CONFLICT 2489 2489

DR PTR; I64034; I64034.  
 KW TIGR; HI1522; -  
 DR Hypochemical protein; Complete proteome.  
 SQ SEQUENCE 623 AA; 67677 MW; 5EE5A8E2818C0060 CRC64;

Query Match 8.4%; Score 83.5; DB 1; Length 623;  
 Best local Similarity 21.7%; Pred. No. 17;  
 Matches 33; Conservative 19; Mismatches 51; Indels 49; Gaps 5;

QY 7 DWCRIMSDVEQKSLMVTGIPAD-----PFEAEIQEVLOETLKSILGRYLLG- 52  
 Db 41 DWFNMVQAEFLSILLDAGIQPKGRNLNQISLAIRKLSGKVDFFSOQ-LQADGYKYIGR 99  
 QY 53 -----KIFRKQENANAVLL-----ELLEDDTVSAIPSE-----VQKGGV 87  
 Db 100 CKSVAELRTIRPTENGORILVDAYYEGSTAGGGEFVADLQDLITDDGGTCFVFPNNGR 159  
 QY 88 WKVIFKPTNQDTFF-----LERNLFLF 110  
 Db 160 WKRLFSSSLQDTHFGVIGGVAADTTNLNAFLD 191

RESULT 15  
 PHYB\_TOBAC  
 ID PHYB\_TOBAC STANDARD; PRT; 1132 AA.  
 AC P29130;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 GN Phytochrome B.  
 GE PHYB.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Nicotiana.  
 OC NCBI\_TaxID=4097;  
 RN [1]  
 RN [2]  
 RN [3]  
 RN [4]  
 RN [5]  
 RN [6]  
 RN [7]  
 RN [8]  
 RN [9]  
 RN [10]  
 RN [11]  
 RN [12]  
 RN [13]  
 RN [14]  
 RN [15]  
 RN [16]  
 RN [17]  
 RN [18]  
 RN [19]  
 RN [20]  
 RN [21]  
 RN [22]  
 RN [23]  
 RN [24]  
 RN [25]  
 RN [26]  
 RN [27]  
 RN [28]  
 RN [29]  
 RN [30]  
 RN [31]  
 RN [32]  
 RN [33]  
 RN [34]  
 RN [35]  
 RN [36]  
 RN [37]  
 RN [38]  
 RN [39]  
 RN [40]  
 RN [41]  
 RN [42]  
 RN [43]  
 RN [44]  
 RN [45]  
 RN [46]  
 RN [47]  
 RN [48]  
 RN [49]  
 RN [50]  
 RN [51]  
 RN [52]  
 RN [53]  
 RN [54]  
 RN [55]  
 RN [56]  
 RN [57]  
 RN [58]  
 RN [59]  
 RN [60]  
 RN [61]  
 RN [62]  
 RN [63]  
 RN [64]  
 RN [65]  
 RN [66]  
 RN [67]  
 RN [68]  
 RN [69]  
 RN [70]  
 RN [71]  
 RN [72]  
 RN [73]  
 RN [74]  
 RN [75]  
 RN [76]  
 RN [77]  
 RN [78]  
 RN [79]  
 RN [80]  
 RN [81]  
 RN [82]  
 RN [83]  
 RN [84]  
 RN [85]  
 RN [86]  
 RN [87]  
 RN [88]  
 RN [89]  
 RN [90]  
 RN [91]  
 RN [92]  
 RN [93]  
 RN [94]  
 RN [95]  
 RN [96]  
 RN [97]  
 RN [98]  
 RN [99]  
 RN [100]  
 RN [101]  
 RN [102]  
 RN [103]  
 RN [104]  
 RN [105]  
 RN [106]  
 RN [107]  
 RN [108]  
 RN [109]  
 RN [110]  
 RN [111]  
 RN [112]  
 RN [113]  
 RN [114]  
 RN [115]  
 RN [116]  
 RN [117]  
 RN [118]  
 RN [119]  
 RN [120]  
 RN [121]  
 RN [122]  
 RN [123]  
 RN [124]  
 RN [125]  
 RN [126]  
 RN [127]  
 RN [128]  
 RN [129]  
 RN [130]  
 RN [131]  
 RN [132]  
 RN [133]  
 RN [134]  
 RN [135]  
 RN [136]  
 RN [137]  
 RN [138]  
 RN [139]  
 RN [140]  
 RN [141]  
 RN [142]  
 RN [143]  
 RN [144]  
 RN [145]  
 RN [146]  
 RN [147]  
 RN [148]  
 RN [149]  
 RN [150]  
 RN [151]  
 RN [152]  
 RN [153]  
 RN [154]  
 RN [155]  
 RN [156]  
 RN [157]  
 RN [158]  
 RN [159]  
 RN [160]  
 RN [161]  
 RN [162]  
 RN [163]  
 RN [164]  
 RN [165]  
 RN [166]  
 RN [167]  
 RN [168]  
 RN [169]  
 RN [170]  
 RN [171]  
 RN [172]  
 RN [173]  
 RN [174]  
 RN [175]  
 RN [176]  
 RN [177]  
 RN [178]  
 RN [179]  
 RN [180]  
 RN [181]  
 RN [182]  
 RN [183]  
 RN [184]  
 RN [185]  
 RN [186]  
 RN [187]  
 RN [188]  
 RN [189]  
 RN [190]  
 RN [191]  
 RN [192]  
 RN [193]  
 RN [194]  
 RN [195]  
 RN [196]  
 RN [197]  
 RN [198]  
 RN [199]  
 RN [200]  
 RN [201]  
 RN [202]  
 RN [203]  
 RN [204]  
 RN [205]  
 RN [206]  
 RN [207]  
 RN [208]  
 RN [209]  
 RN [210]  
 RN [211]  
 RN [212]  
 RN [213]  
 RN [214]  
 RN [215]  
 RN [216]  
 RN [217]  
 RN [218]  
 RN [219]  
 RN [220]  
 RN [221]  
 RN [222]  
 RN [223]  
 RN [224]  
 RN [225]  
 RN [226]  
 RN [227]  
 RN [228]  
 RN [229]  
 RN [230]  
 RN [231]  
 RN [232]  
 RN [233]  
 RN [234]  
 RN [235]  
 RN [236]  
 RN [237]  
 RN [238]  
 RN [239]  
 RN [240]  
 RN [241]  
 RN [242]  
 RN [243]  
 RN [244]  
 RN [245]  
 RN [246]  
 RN [247]  
 RN [248]  
 RN [249]  
 RN [250]  
 RN [251]  
 RN [252]  
 RN [253]  
 RN [254]  
 RN [255]  
 RN [256]  
 RN [257]  
 RN [258]  
 RN [259]  
 RN [260]  
 RN [261]  
 RN [262]  
 RN [263]  
 RN [264]  
 RN [265]  
 RN [266]  
 RN [267]  
 RN [268]  
 RN [269]  
 RN [270]  
 RN [271]  
 RN [272]  
 RN [273]  
 RN [274]  
 RN [275]  
 RN [276]  
 RN [277]  
 RN [278]  
 RN [279]  
 RN [280]  
 RN [281]  
 RN [282]  
 RN [283]  
 RN [284]  
 RN [285]  
 RN [286]  
 RN [287]  
 RN [288]  
 RN [289]  
 RN [290]  
 RN [291]  
 RN [292]  
 RN [293]  
 RN [294]  
 RN [295]  
 RN [296]  
 RN [297]  
 RN [298]  
 RN [299]  
 RN [300]  
 RN [301]  
 RN [302]  
 RN [303]  
 RN [304]  
 RN [305]  
 RN [306]  
 RN [307]  
 RN [308]  
 RN [309]  
 RN [310]  
 RN [311]  
 RN [312]  
 RN [313]  
 RN [314]  
 RN [315]  
 RN [316]  
 RN [317]  
 RN [318]  
 RN [319]  
 RN [320]  
 RN [321]  
 RN [322]  
 RN [323]  
 RN [324]  
 RN [325]  
 RN [326]  
 RN [327]  
 RN [328]  
 RN [329]  
 RN [330]  
 RN [331]  
 RN [332]  
 RN [333]  
 RN [334]  
 RN [335]  
 RN [336]  
 RN [337]  
 RN [338]  
 RN [339]  
 RN [340]  
 RN [341]  
 RN [342]  
 RN [343]  
 RN [344]  
 RN [345]  
 RN [346]  
 RN [347]  
 RN [348]  
 RN [349]  
 RN [350]  
 RN [351]  
 RN [352]  
 RN [353]  
 RN [354]  
 RN [355

[illegible]

Search completed: September 21, 2004, 13:32:37  
Job time : 6.31599 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 14.4981 Seconds  
(without alignments)  
694.369 Million cell updates/sec

Title: US-10-037-860-7  
Perfect score: 996  
Sequence: 1 PLALLEDWCRIMSVDEQKSL.....EBESFEVWLEQATEIVKEMP 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	996	100.0	195	4	US-09-189-527-7
2	443.5	44.5	462	4	US-09-189-527-13
3	441	44.3	329	4	US-09-189-527-4
4	87.5	8.8	311	2	US-08-318-837-9
5	82.5	8.3	547	4	US-09-134-000C-5974
6	81.5	8.2	580	4	US-09-328-352-7656
7	80.5	8.1	136	4	US-09-252-991A-31474
8	80.5	8.1	270	2	US-08-852-743-5
9	80.5	8.1	270	3	US-03-185-370-5
10	80.5	8.1	487	3	US-08-712-709-8
11	80.5	8.1	487	3	US-09-111-444-8
12	80.5	8.1	487	3	US-09-541-228-8
13	80	8.0	375	4	US-09-328-352-7783
14	79	7.9	258	4	US-09-252-991A-24184
15	78.5	7.9	316	1	US-08-403-634-4
16	78.5	7.9	316	1	US-08-403-634-31
17	78.5	7.9	316	3	US-08-913-441B-4
18	78.5	7.9	316	3	US-08-913-441B-31
19	78.5	7.9	445	3	US-09-457-046B-52
20	78.5	7.9	745	1	US-08-136-277-2
21	78.5	7.9	745	2	US-08-479-403-2
22	78.5	7.9	745	3	US-08-835-734-2
23	77.5	7.8	1657	1	US-08-287-959-1
24	77.5	7.8	1805	1	US-07-853-913-2
25	77	7.7	473	4	US-09-252-991A-29636
26	77	7.7	542	4	US-09-489-847-323
27	76	7.6	248	4	US-09-489-039A-12849

28	76	7.6	759	4	US-09-252-991A-30106
29	75.5	7.6	475	4	US-09-370-838-193
30	75	7.5	419	4	US-09-134-000C-4859
31	75	7.5	1048	4	US-08-887-534A-85
32	75	7.5	1048	4	US-09-527-431-85
33	75	7.5	1087	4	US-09-914-259-12
34	74.5	7.5	430	4	US-09-489-039A-7776
35	74.5	7.5	2108	4	US-09-252-991A-31502
36	74	7.4	227	4	US-09-252-991A-25961
37	74	7.4	388	4	US-09-489-039A-9777
38	74	7.4	407	2	US-08-934-846-4
39	74	7.4	407	3	US-09-238-557-4
40	74	7.4	418	2	US-08-934-846-2
41	74	7.4	418	3	US-09-238-557-2
42	74	7.4	521	3	US-08-948-564-14
43	74	7.4	548	4	US-09-328-352-7955
44	73.5	7.4	227	2	US-08-698-407-1
45	73.5	7.4	227	3	US-09-195-855-1

ALIGNMENTS

RESULT 1  
US-09-189-527-7  
; Sequence 7, Application US/09189527A  
; Patent No. 6387639  
; GENERAL INFORMATION:  
; APPLICANT: Jerome B. Posner  
; APPLICANT: Josep O. Dalmau  
; APPLICANT: Myrna R. Rosenfeld  
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma  
; TITLE OF INVENTION: Antibodies  
; FILE REFERENCE: SLK98-01  
; CURRENT APPLICATION NUMBER: US/09/189,527A  
; CURRENT FILING DATE: 1998-11-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-189-527-7

Query Match	100.0%;	Score	996;	DB	4;	Length	195;
Best Local Similarity	100.0%;	Pred. No.	1.4e-98;				
Matches	195;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	PLALLEDWCRIMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQEN	60				
Db	1	PLALLEDWCRIMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQEN	60				
Qy	61	ANAVLLELLEDTVSALPSEVQGGYKVIKFTPNQDTFELRLNLFLEKEGQTVSGMF	120				
Db	61	ANAVLLELLEDTVSALPSEVQGGYKVIKFTPNQDTFELRLNLFLEKEGQTVSGMF	120				
Qy	121	RALGQEALS PATVPCISPELLAHLLGQMAHAPOLPLPMRYRKLRFVSGSAVPAPEESF	180				
Db	121	RALGQEALS PATVPCISPELLAHLLGQMAHAPOLPLPMRYRKLRFVSGSAVPAPEESF	180				
Qy	181	EWLEQATEIVKEMP	195				
Db	181	EWLEQATEIVKEMP	195				

RESULT 2  
US-09-189-527-13  
; Sequence 13, Application US/09189527A  
; Patent No. 6387639  
; GENERAL INFORMATION:  
; APPLICANT: Jerome B. Posner  
; APPLICANT: Josep O. Dalmau  
; APPLICANT: Myrna R. Rosenfeld

;; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma  
;; TITLE OF INVENTION: Antibodies  
;; FILE REFERENCE: SLK98-01  
;; CURRENT APPLICATION NUMBER: US/09/189,527A  
;; CURRENT FILING DATE: 1998-11-10  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 13  
;; LENGTH: 462  
;; TYPE: PRT  
;; ORGANISM: homo sapiens  
US-09-189-527-13

Query Match 44.5%; Score 443.5; DB 4; Length 462;  
Best Local Similarity 46.8%; Pred. No. 6e-39; Indels 3; Gaps 2;  
Matches 89; Conservative 34; Mismatches 64; Indels 3; Gaps 2;  
QY 6 EDWCRIMSDVEQKSLMTVGIPADPEAEIQEVLQETLKSIGRYRLGKIFRKQENANAVL 65  
DB 1 QDWCRGHEHNRRCMLLIGPEDCGEDEFEEITLQEACRHLGRYRVIGRMFRRENAQAIL 60  
QY 66 LLELLEDDTVSAIPSEVQGGVGMKVIKFTPNQDTEFLERLNLFLKEGQTVSGMFRALQ 125  
DB 61 LELAQDIDVALLPREIPGKGGPWEVIVKPRNSDGEFLNRLNRFLEBERRTVSDMNRVLGS 120  
QY 126 EALSPATVPCISPELLAHLHLLQMAHAPOPLL-PMRYRKLRFVSGSAVPAPPEESFEVWL 184  
DB 121 DYNCSAPRVTISPEFWT--WAOTLGAAVQPLLEQMLYRELRFVSGNTISIPGALAFDAWL 178  
QY 185 EQATEIVKWE 194  
DB 179 EHTTEMLQMW 188

RESULT 3  
US-09-189-527-4  
; Sequence 4, Application US/09189527A  
; Patent No. 6387639  
; GENERAL INFORMATION:  
; APPLICANT: Jerome B. Posner  
; APPLICANT: Josep O. Dalmau  
; APPLICANT: Myrna R. Rosenfeld  
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma  
; TITLE OF INVENTION: Antibodies  
; FILE REFERENCE: SLK98-01  
; CURRENT APPLICATION NUMBER: US/09/189,527A  
; CURRENT FILING DATE: 1998-11-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 329  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-189-527-4

Query Match 44.3%; Score 441; DB 4; Length 329;  
Best Local Similarity 46.4%; Pred. No. 6.8e-39; Indels 6; Gaps 4;  
Matches 90; Conservative 37; Mismatches 61; Indels 6; Gaps 4;  
QY 2 LALLEDDCRIMSDVEQKSLMTVGIPADPEAEIQEVLQETLKSIGRYRLGKIFRKQENA 61  
DB 3 MTLLEDDCRGMDVNSQRTLLVWGPVNCDEAEIETLQAAPOVS-YRMLGRMFWRENA 61  
QY 62 NAVLELLEDDTVSAIPSEVQGGVGMKVIKFTPNQDTEFLERLNLFLKEGQTVSGMFR 121  
DB 62 KAALLELTGAVDYAIPREMPGKGGVGMKVIKFTPNQDTEFLERLNLFLAREGWTVDVAR 121  
QY 122 ALQGEALSPATVPCISPELLAHLHLLQMAHAPOPLL-PMRYRKLRFVSGSAVPAPPEESF 180  
DB 122 VLGFQ--NPTTP--GPMPAPMLNILDNIQPLVESIWTYKRLTLFSGKGHPRAWGNF 177  
QY 181 EWLLEQATEIVKWE 194

Db 178 DPMLEHTNEVLEEW 191

RESULT 4

US-08-318-837-9  
; Sequence 9, Application US/08318837  
; Patent No. 5981277  
; GENERAL INFORMATION:  
; APPLICANT: FRANSEN, LUCIA; DEVOS, KATHLEEN; VAN DE VOORDE,  
; APPLICANT: ANDRE; VAN HEUVERSWYN, HUGO  
; TITLE OF INVENTION: NEW POLYPEPTIDES AND PEPTIDES, NUCLEIC ACID  
; TITLE OF INVENTION: CODING FOR THEM, AND THEIR USE IN THE FIELD OF TUMOR THERAPY  
; TITLE OF INVENTION: IMMUNOLOGY  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: BIERMAN AND MUSERLIAN  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,837  
; FILING DATE: 13-OCT-1994  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP 93/01022  
; FILING DATE: 28-APR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 92,401,231.3  
; FILING DATE: 30-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHARLES A. MUSERLIAN  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 410.007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 661-8000  
; TELEFAX: (212) 661-8002  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 311 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-318-837-9

Query Match 8.8%; Score 87.5; DB 2; Length 311;  
Best Local Similarity 27.1%; Pred. No. 0.4;  
Matches 45; Conservative 21; Mismatches 59; Indels 41; Gaps 11;  
QY 53 KIFRKQENANAVL-----LLELLEDDTVSAIPSEVQ-----GKGWVWVIFKFTPNQDTEFE 103  
DB 109 KPFFDSGGANLYLEKTKTELRLV-RDIRGPGVQVQCFSLQGGGLF--VEATPOODIS--R 163  
QY 104 RLNLFLKEGQTVSGMFRALQGEALSPATVPCISPELLAHLG-----QAMAHAPQ 154  
DB 164 RTTGFG---QVELMSGQ-RGLDLHVLSAPCRPCSDTEVLLAICTSDTFVVRGFIEDVTHYPE 219  
QY 155 PLLPMRY-----RKLRFVSGSAVPAPPEESFEVWLEQATEIVK 192  
DB 220 QQVSVIYLRNRLHRQKSRVQ-----FAPEDSGH--WLGHVTTLLQ 259

RESULT 5  
US-09-134-000C-5974  
; Sequence 5974, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:

```

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5974
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-5974

```

Query Match	8.3%;	Score 82.5;	DB 4;	Length 547;
Best Local Similarity	22.0%;	Pred. No. 3.2;		
Matches	44;	Conservative 38;	Mismatches 65;	Indels 53; Gaps 9;

  

QY	5	LEDW	--CRIMSDVEQKSLMVTGI--PAD----	FEAAE-----	LOEVQETLKS	GRY	48
Db	310	LELWKVRLNSDYQLAIVGYTKENETHYRQQREGQLIFQWLKEQPEILPDVAF					369
QY	49	RLGK	-----IPKQNNANVLELLEDDTVSAIPSEVQ--GKGGMVKVIFKPTNQDTEFL				102
Db	370	KLQNQKSLILIFOSKNDHLMILQNLAERLQQLPITIRPALGNAYENLEDLNSYEAS					429
QY	103	ERLNLFLEKEGQTVSGMFRALGOEALSPATVPCISPELLAHL---	LG-----				147
Db	430	STLEASL-----	HAQKPAIVQLFHPKGLAGLFEKIGTDEVEYFCQQQLK				473
QY	148	AMAHAPQPLLPMYRKLRFV					167
Db	474	ELAYTEPTLQELRKLTKVF					493

```

RESULT 6
US-09-328-352-7656
; Sequence 7656, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Brston et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7656
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7656

```

	Query Match	8.2%;	Score 81.5;	DB 4;	Length 580;
	Best Local Similarity	25.2%;	Pred. No. 4.4;		
	Matches	36;	Conservative	22;	Mismatches 58; Indels 27; Gaps 5;
QY	30	EEAEIQEVLOETLKS	LR-----YRL-LGKIFRKQNNANAVLLELEDDTVS	75	
Db	248	EQGTAEQVLEQPKDVYTRALLYCRPQMSQRPYLPVTSDFMRQE--NNILVE--QSFDSV	303		
QY	76	AIPSEVQKGGVKNVIEKPTNQDTEFLRLNLFLEKSGQTVSGM-----FRALQGE	126		
Db	304	EIPERKEGLNGDEQIILEVKDLKSFYSRKGLFKGEFFQAVKGVSKLAKGKTGLGVCS	363		
QY	127	ALSPATVPCTSPPELLAHLGQAM	149		
Db	364	GSCKTTVGILLMRLHQASGGQAL	386		

## RESULT 7

```

US-09-252-991A-31474
; Sequence 31474, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31474
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31474

```

Query Match	8.1%	Score 80.5;	DB 4;	Length 136;
Best Local Similarity	21.6%;	Pred. No.	0.67;	
Matches	41; Conservative	24; Mismatches	48; Indels	77; Gaps
9;				
QY	17	OKSLMTGTP-----ADFEAEIQEVLQETLSIGRYVLLGKTFRKQENANAVLLELLED	71	
Db	9	QHIFITGVKRNKSLAEFREAE     :     :     :		
QY	72	TDVSAIPSEVQGGVGWVKVIFKTPNQDTEFLERLNLFLEXEGOTVSGMFRALGOEALSPA	131	
Db	47	SSL-----KQELEPKDKLQALMDKYGMTLHNI IAILDPKA--PV	83	
QY	132	TVPICISPELLAHLGQAMAHAPOLPMPRVKRLRPF-----SGSAAVPAP---EESFSFYWLE	185	
Db	84	TV-----SAAPO-----RRARALKVYNPNNGVEVETKGGNHKVLKAWKE	123	
QY	186	Q-ATEIVKEW	194	
Db	124	QYGSETVESW	133	

```

RESULT 8
US-08-852-743-5
; Sequence 5, Application US/08852743
; Patent No. 5830699
; GENERAL INFORMATION:
; APPLICANT: Force, Thomas
; APPLICANT: Kyriakis, John M.
; APPLICANT: Pombo, Celia M.
; APPLICANT: Bonventre, Joseph
; TITLE OF INVENTION: SOK-1 AND METHODS OF USE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,743
; FILING DATE: 7-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,774
; FILING DATE: 7-MAY-1996
; ATTORNEY/AGENT INFORMATION:

```





Query Match 7.9%; Score 79; DB 4; Length 258;  
Best Local Similarity 21.5%; Pred. No. 2.5;  
Matches 41; Conservative 33; Mismatches 77; Indels 40; Gaps 8;  
QY 17 OKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKOENANAVLL-----E 67  
DB 8 RKLITLITGASSGFGQALAREALD-----AGHRVVGTV--RSEARSALAEVAPQOAFGR 59  
QY 68 LIEDTDVSAIP---SEVOGKGWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALG 124  
DB 60 LLDVTDLAARPTVAIERDIGPLDVNSAGYGHGIEESPLAEMRRQFEVNLF---G 116  
QY 125 QEALSPATVPCISPELLAHLGQMAHAPOPLPMRYRKLRFVSGSAVAPAEESFEVWL 184  
DB 117 AVAMIQAVLPYMRRRRRGHILNRLHHRHARHLLRQ-----QVRPGRGFKP-----W- 165  
QY 185 EQATEIVKEMP 195  
DB 166 -----ARRWP 170  
RESULT 15  
US-08-403-634-4  
; Sequence 4, Application US/08403634  
; Patent No. 5674748  
; GENERAL INFORMATION:  
; APPLICANT: Giordano, Antonio  
; TITLE OF INVENTION: NOVEL HUMAN CYCLIN-DEPENDENT  
; TITLE OF INVENTION: KINASE-LIKE PROTEINS AND METHODS  
; TITLE OF INVENTION: OF USING THE SAME  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &  
; ADDRESS: No. 5674748ris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,634  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/208,575  
; FILING DATE: 08-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TCU-1482  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 316 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-403-634-4  
Query Match 7.9%; Score 78.5; DB 1; Length 316;  
Best Local Similarity 25.5%; Pred. No. 3.8;  
Matches 40; Conservative 23; Mismatches 65; Indels 29; Gaps 6;  
QY 5 LEDMCRIMSVDEQK-SLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKOENANA 63  
DB 83 LESIELVMGYCEQDLASILENNPTFFSEAOVKCIVLQVLRGL-QYLHRRNFIIHRLKVEN 141

Query Match 8.0%; Score 80; DB 4; Length 375;  
Best Local Similarity 21.2%; Pred. No. 3.4; Indels 72; Gaps 10;  
Matches 42; Conservative 35; Mismatches 49; Indels 72; Gaps 10;  
QY 14 VDEOKSLMVTGIPADFE-----EAEIQEVL---QETLKSIGRY--RLLGKIFRQEN 60  
DB 100 IDNRQALLIATPNDPESKLNQSAPIQIVDGRNSSTAWAGSYLNKIIGQFNQOKFN 159  
QY 61 ANAVLLELLEDDTVSAIPSEVOGKGWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMF 120  
DB 160 -----SALPISLETR--TW-----YNPQESRW----- 180  
QY 121 RALGQEALSPATVPCISPELLAHLGQMAHAPO-----PLIPMYRKLRFVSGSAVP 173  
DB 181 -----SIMPALIAALS--MMQTLLSALSAREEQGTDPQLLVTPYPLQIMIGRALP 232  
QY 174 APEESFEVWLEQATEIV 191  
DB 233 -----PIFVGLMQSTIIL 245  
RESULT 14  
US-09-252-991A-24184  
; Sequence 24184, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24184  
; LENGTH: 258  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24184

US-09-328-352-7783  
; Sequence 7783, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7783  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7783

DB 149 DIKAGNILL-----NTEGHAKLADFGVAGQLTDTWAKEN-----TV 184  
QY 117 SGMFRALGQEALSPATVPCISPELLAHLGQMAHAPO---LLPMRYRKLRFVSGSAVP 173  
DB 185 IGTPWMAPEVTEIGYNCVADIWSLIGITAIEMAGKRPYADIHPMR-----AIFMPTNP 240  
QY 174 APEESFEVWLEQATEIVKE 193  
DB 241 PPTFRKPELWSDNFTDFVKQ 260

```
Qy      64 VLLELEDTVSALPSEVQKGGVMKVIFX--TENQDTEFLERLNLFEKEGQTVSGMFR 121
Db      142 LLM-----TDKGCVKTADFGLARAYGVVKPMTPKVVTLWYRAPELLGTTTQTTSIDMW 196
Qy      122 ALGOEALSPATVPCISPELLAHLIGQAMAHAPQPLLP 158
Db      197 AVG-----CILAEHLAH-----RPLLP 213
```

Search completed: September 21, 2004, 13:36:17  
Job time : 15.4981 secs

Blank

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 21.3916 Seconds  
(without alignments)  
1272.567 Million cell updates/sec

Title: US-10-037-860-11

Perfect score: 1462

Sequence: 1 VQKGGVMKVIKTPNQDTE.....SIEPEERDGYGRWNHEGDD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	7.7	554	2 JMW0094	neurofilament prot
2	112	7.7	2442	2 T08621	centrosome associa
3	110.5	7.6	544	2 B44841	low molecularweig
4	105.5	7.2	659	2 S11736	resistance protein
5	105.5	7.2	659	2 S11737	resistance protein
6	103.5	7.1	991	2 C98232	hypothetical prote
7	103	7.0	374	2 A83483	ribonuclease D PAI
8	103	7.0	544	2 S07144	neurofilament trip
9	103	7.0	803	2 T39530	hypothetical prote
10	103	7.0	1840	2 T29091	transitin - chicke
11	102.5	7.0	614	2 AB2304	two-component hybr
12	102.5	7.0	879	2 AH1262	alanyl-tRNA synthe
13	101.5	6.9	520	2 AD2616	chromosomal replic
14	101.5	6.9	529	2 D97398	dnAA protein (L254
15	100.5	6.9	548	1 QFPGU	neurofilament trip
16	100.5	6.9	558	2 D70449	conserved hypothet
17	100.3	6.9	879	2 AB1625	alanyl-tRNA synthe
18	100	6.8	733	2 F82965	hypothetical prote
19	99.5	6.8	315	2 T26186	hypothetical prote
20	99.5	6.8	543	1 QFMSL	neurofilament trip
21	99.5	6.8	2712	2 T05113	hypothetical prote
22	99	6.8	399	2 G72253	RNA polymerase sig
23	98.5	6.7	1175	2 C35815	myosin heavy chain
24	98.5	6.7	1175	2 D35815	myosin heavy chain
25	98.5	6.7	1201	2 A35815	myosin heavy chain
26	98.5	6.7	1201	2 B35815	myosin heavy chain
27	98.5	6.7	1804	2 T34318	nestin - golden ha
28	98.5	6.7	2385	2 A32491	myosin heavy chain
29	98.5	6.7	2411	2 B32491	myosin heavy chain

#### ALIGNMENTS

##### RESULT 1

JMW0094

neurofilament protein-L - bovine

N:Alternate names: NF-L

C:Species: Bos primigenius taurus (cattle)

C>Date: 03-Jul-1998 #sequence\_revision 10-Jul-1998 #text\_change 07-May-1999

C:Accession: JMW0094

R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kaibuchi, K.; Inagaki, M.

A:Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated kinase

A:Reference number: JMW0094; MUID:98238650; PMID:9571164

A:Accession: JMW0094

A:Molecule type: protein

A:Residues: 1-554 <HAS>

C:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly,

C:Superfamily: cytoskeletal keratin

Query Match 7.7%; Score 112; DB 2; Length 554;

Best Local Similarity 24.2%; Pred. No. 1.6;

Matches 55; Conservative 30; Mismatches 82; Indels 60; Gaps 10;

QY	100	FEWLEQA---TEIVKEVPTVEAEKKWL-----AESLRG--PALD--LMHIVQADNPS	146
Db	284	FTVLTESAAKNVTDVRAAKDVESESRLLKAKTLEIACRGMNEALEKQLQLEDKQVAD	343
QY	147	ISVEECLEAFKQVFGSLSESRRTAQVRYLKTQVE-----EKEK-----	183
Db	344	ISAMQ---DTINKLENEIRTKSEMARYLKEYQDLLNVMALDIIAAYRKLLEGEETRLS	401
QY	184	-----VSAYVLRLETLRKA-----VEKRAIPRRADQVRLEQVMAGATLNQM	226
Db	402	FTSVGSLTTGTQSSQVFGSAVGLQTSSVYLSARSFPSTVYTSVHVEEQIEVETIEA-	460
QY	227	LWCLRLRLKQGGPPSPFLELMKVIREEEEEAASFENISEEPBERDG	273
Db	461	--AKAEAKDE--PPSEGEAEKEKEAEAEAEAEAEAEAEAEAE	503

##### RESULT 2

T08621

centrosome associated protein CEP250 - human

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: T08621

R:Mack, G.J.; Rees, J.; Sandblom, O.; Balczon, R.; Fritzler, M.J.; Rattner, J.B.

A:Title: Autoantibodies to a group of centrosomal proteins in human autoimmune sera react

A:Reference number: Z16462; MUID:98165428; PMID:9506584

A:Accession: T08621

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-2442 <MAC>

A;Cross-references: EMBL:AF022655; NID:G2832236; PIDN:AAC06349.1; PID:G2832237  
A;Experimental source: cell line HeLa

Query Match 7.7%; Score 112; DB 2; Length 2442;  
Best Local Similarity 23.4%; Pred. No. 10;  
Matches 76; Conservative 47; Mismatches 94; Indels 108; Gaps 17;  
QY 22 LERLNLFLKEGQTVSGMFRALGQGVSPATVPCISPELLAHLGQMAHAPOPLPM-- 79  
Db 1458 LELLSLDLKRNQEVN-----LQEQIQELE-KCRS--VLEH-----LPMVA 1496  
QY 80 --RYKRLRVFSSAVPAPEEESFEVWLEQATEIVKEWPVTEAEKKRWLAESLRGPALDI- 136  
Db 1497 QEREQKLTIV-QREQIREPKD-----RETQNVLEHQLLEKDDQMIESRGQVQDLK 1549  
QY 137 MHIVQADNPISIVE-----ECLE-AFKQVFGSLSRRTA----- 169  
Db 1550 KQVLTLECLALEENHNHKEWCOOKLIKELEGQRTQRTVALTHLTLDLERSQELQAQSS 1609  
QY 170 QVRYLKY-----QEEGKVSAYVLETLTKA-----VEKR----- 202  
Db 1610 QIHDLESHSTVLAELQEQDQVKSQREIQEELQROKHEHTQDLRRDQELMLQKRIQV 1669  
QY 203 -----AIPRIADQVRLEQVMAG---ATLNQMLWCLRELKDKQGGPPS-----FILE 245  
Db 1670 LEDQRTQTKILEEDLEQKLSRERGRLTTQRL-----MQEAEKGKPSKAQSGSLE 1725  
QY 246 LMKVIREEEERASFNESIEEPEE 270  
Db 1726 HMKLILRDKEVECCQEHILEQE 1750

RESULT 3  
B44841  
low molecular weight neurofilament protein XNF-L - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 31-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 12-Apr-1995  
C;Accession: B44841  
J;Chatnas, L.R.; Szaro, B.G.; Gainer, H.  
R;Neurosci. 12, 3010-3024, 1992  
A;Title: Identification and developmental expression of a novel low molecular weight neu  
A;Reference number: A44841; MUID:92356194; PMID:1494944  
A;Accession: B44841  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-544 <CHA>  
A;Experimental source: brain  
A;Note: sequence inconsistent with the nucleotide translation  
A;Note: sequence extracted from NCBI backbone (NCBIN:110225, NCBIP:110226)  
C;Superfamily: cytoskeletal keratin

Query Match 7.6%; Score 110.5; DB 2; Length 544;  
Best Local Similarity 22.3%; Pred. No. 2;  
Matches 59; Conservative 30; Mismatches 83; Indels 93; Gaps 10;  
QY 80 RYKRLRVFSSAVPAPEEESFEVWLE-----QATEIVKEWPVTEAEKKRWLA-- 126  
Db 258 QYKYL-----AAKMQSAEENFKGRFTVLTQSAARNTDAVRAAKDENSESRMLSAK 309  
QY 127 -----ESURGPALDLMHIVQ--ADNPISIVBECLEAFKQVFGSLSRRTAQVRYLKYQE- 179  
Db 310 GLEIEACRGVNEALQRIQIELEDKQSGEAGMQDAINKLEBELNTRKSEMARYLKEYQDL 369  
QY 180 -----EBKVV-----SAYVLRLETLRKA 198  
Db 370 LNVKMLDIEIAAYRKLEGBETRLFSFGVGAITSYGTQSAVFGRSAYLSQSSYM--- 426  
QY 199 VEKALPRIADQVRLEQVMAGATLNQMLWCLRELKDKQGGPPSFLMLKVIREEESEA 258  
Db 427 -TSRAFFPYSHVQEEQLDIETIES---SRAEAKAEAP-----EEEEEA 471  
QY 259 SFENESIEEPEERDGYGRWNHGGDD 283

Db 472 EEEGGGSEAE-----GEEGEE 490  
RESULT 4  
S11736  
resistance protein Mx2, interferon-regulated - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 19-May-2000  
C;Accession: S11736  
R;Meier, E.; Kunz, G.; Haller, O.; Arnheiter, H.  
submitted to the EMBL Data Library, April 1990  
A;Description: Unexpected activity of rat Mx proteins against a Rhabdovirus.  
A;Reference number: S11735  
A;Accession: S11736  
A;Molecule type: mRNA  
A;Residues: 1-659 <EMB>  
A;Cross-references: EMBL:X52712; NID:G56722; PIDN:CAA36936.1; PID:G56723  
C;Superfamily: dynamin-related protein VPS1

Query Match 7.2%; Score 105.5; DB 2; Length 659;  
Best Local Similarity 21.5%; Pred. No. 5.8;  
Matches 64; Conservative 51; Mismatches 84; Indels 99; Gaps 16;  
QY 20 EFLERLNLFLKEGQTVSGMFRALGQGVSPATVPCISP-----ELLAHLGQMAHAPOP 75  
Db 292 EALQEQVFFKEHPQ-----FRALLEDG--KATVPCLAERLTWELISHICKS----- 336  
QY 76 LLPMYRKLRVFSGSAVPAPEEESFEVWLEQATEIVKEWPVTEAEKKRWLAESLRGPALD 135  
Db 337 -LPLENQIK-----ESHQSTSEELQYKADIPEDENEXTLFLIEKINAFNQD 383  
QY 136 LMHIVQADNPISIVEEC---LEAFKQVF-GSLESRTAQ-----VRYLKY-----QEEGK 183  
Db 384 ITAIVEGEE-IVREKECRLFTKLKEFFLWSEIERNFQKSGDALYKEVYTFMFQYKRE 442  
QY 184 VSAVY-----LRLETLRKAVE-----KRAIPR 206  
Db 443 LQGFVNYKTFENIRROIKTLEPAMEMLHKVTEIVRAAFTTVSEKNFSEFFNLHRTKS 502  
QY 207 RIADQVRLEO-----VMAGATLNQMLWCLRELKDKQGGPPSFLMLKVIREEESEE 257  
Db 503 KLED-IRLQETAEAKSIRLHFQMEQIIVC-----QDQ-----IYRKALQKVREEEAE 550

RESULT 5  
S11737  
resistance protein Mx3, interferon-regulated - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 17-Mar-2000  
C;Accession: S11737  
R;Meier, E.; Kunz, G.; Haller, O.; Arnheiter, H.  
submitted to the EMBL Data Library, April 1990  
A;Description: Unexpected activity of rat Mx proteins against a Rhabdovirus.  
A;Reference number: S11735  
A;Accession: S11737  
A;Molecule type: mRNA  
A;Residues: 1-659 <EMB>  
A;Cross-references: EMBL:X52713; NID:G56724; PIDN:CAA36937.1; PID:G56725  
C;Superfamily: dynamin-related protein VPS1

Query Match 7.2%; Score 105.5; DB 2; Length 659;  
Best Local Similarity 21.5%; Pred. No. 5.8;  
Matches 64; Conservative 51; Mismatches 84; Indels 99; Gaps 16;  
QY 20 EFLERLNLFLKEGQTVSGMFRALGQGVSPATVPCISP-----ELLAHLGQMAHAPOP 75  
Db 292 EALQEQVFFKEHPQ-----FRALLEDG--KATVPCLAERLTWELISHICKS----- 336  
QY 76 LLPMYRKLRVFSGSAVPAPEEESFEVWLEQATEIVKEWPVTEAEKKRWLAESLRGPALD 135  
Db 337 -LPLENQIK-----ESHQSTSEELQYKADIPEDENEXTLFLIEKINAFNQD 383  
QY 136 LMHIVQADNPISIVEEC---LEAFKQVF-GSLESRTAQ-----VRYLKY-----QEEGK 183

Db	384	ITAI	VEGEE	-IVRE	CECLF	YKLR	KEFFL	WSEIE	ERNFQ	SGSDALY	KEVYTF	FMQYRGRE	442
QY	184	VSAYV	-----	LR	ETLLR	KAVE	-----						
Db	443	LPGF	VNYKTF	ENIR	RQIKT	LEEP	AMEML	HKVTEI	VRAA	FTTVSE	KNFSEFF	NLHRTTKS	502
QY	207	RIADQ	VRLEQ	-----	VM	AGATL	NOMLW	CHREL	KCOG	PPPPF	LEMLK	VIREEEE	257
Db	503	KUED	-IRLEQ	ETEAK	ILRH	PHFME	QIYIC	-----	ODQ	---	YRKALQ	KVREEAAE	550

RESULT 6  
C98232  
Hypothetical protein AGR\_L1617 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
C/Species: Agrobacterium tumefaciens  
C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
C/Accession: C98232  
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurolo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A/Reference number: A97359; MUID:21608551; PMID:11743194  
A/Accession: C98232  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-991 <XUR>  
A/Cross-references: GB:AE007870; PIDN:AAK89381.1; PID:g15159232; GSPDB:GN00170  
C/Genetics:  
A/Gene: AGR\_L1617  
A/Map position: linear chromosome

Query Match	7.1%;	Score 103.5;	DB 2;	Length 991;
Best Local Similarity	24.6%;	Pred. No. 14;		
Matches 75;	Conservative 34;	Mismatches 101;	Indels 95;	Gaps 15;

QY		41	RALGOEVSATVPCISPE-----LLAHLGCAHAHAPQLPMRYKLRVFSGSAV	92
Dd		81	RALG-----GTVGFIWPEIWDNWSRILEAGFRCEVNAYRDQPWTLMRH	123
QY	:::	93	PAPEERSFEVWLEQATEIVKE-----WPVTEAEKKWLAEBSIRGPALDLMHIVOA--	142
Dd		124	GVPBEVIFDLFY---TPINYEGTVGVLCVTULENTDKVRALEAESRMELSRLTNALP	180
QY	:::	143	-----DNPSI-----SVECELEAF-----KQVFGS-----LESRRTAQVRVLKYQE	179
Dd	:::	181	ILVGYDRDVVYFANDGYLEWFGRRAEEVIGRSVPDIVGAAPFARRTYLDRAL----	235
QY		180	EGEK-VSAYVLRLETLLRKAVEKRAIPERRAD-----OVRELOVM-----	218
Dd		236	AGEKIUSDVIRPEDGSLRAELISYPVRYLSDSGIDGIYVLIIDIERKKESEQBILTSNN	295
QY	: ::	219	----ACATLNMLWCRIRELKDGGPPSPFLMKVIREEEEEEPASFENESIEEPEDRG-	273
Dd	: ::	296	RFAAAVEAHGVLTNSADGRMERGEOPAWTAM---TGQTPEEYQDFGWADVHPEDRQGS	352
QY		274	YGRWN	278
Dd		353	VDSWN	357

RESULT 7  
A83483  
C:ribonuclease D PA1294 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: A83483  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A:Reference number: A82950; PMID:20437337; PMID:10984043

A;Accession: A83483  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-374 <STO>  
A;Cross-references: GB:AE0045  
A;Experimental source: strain  
C;Genetics:  
A;Gene: rmd; PAL294  
C;Superfamily: ribonuclease D

Query Match 7.0%; Score 103; DB 2; Length 374;  
Best Local Similarity 23.1%; Pred. No. 4.3;  
Matches 65; Conservative 34; Mismatches 90; Indels

QY	23	ERINLFLBKEGTVSGMFRALQGQGVSPATVPCISPPELLAHLGLQAMAHAPQPLLMRYR	82
Db	87	EDLEVFLRLTGS-----LPVPLFDTQLAAAYLG--MAHS-----MGVS	122
QY	83	KLRVFGSAPVAPREESFEVWLEQATEIVKEWPVTEAEKKWLAESIRGPALDLMHIVQA	142
Db	123	KL-VKEVLIDILPKDSTRDWLQR-----PLTEMQ-----MRYAADVOQLHACV	165
QY	143	-----DNPSISVEFC-----LEAFQV-FG-SLESRRTAQVRYLKTY	177
Db	166	YIALDTRLSEKRAWLLEDGAELVANLCRESDPREAYREVKLGRWRPQOLAVLRELCAW	225
QY	178	OBEQEKV-----SAVVLRETLRLKAVEKRAIPRIADQVRLQ-----V	217
Db	226	REEQARLNRPNRHVLBERTLWPLA---RLLPKNKTDLAAIEDMHPTVTRQGDFLIELI	282
QY	218	MAGATLNMQLWCRLRELKUDQGPPSPFLMLMKVIREBEEEA	258
Db	283	ACAARLPQSEW---PEALPFLPPEVTPLIKSLRAIGOREA	320

## RESULT 8

S07144  
neurofilament triplet L protein - human  
N/Alternate names: neurofilament light polypeptide (68k)  
N/Contains: Glu-50 brain peptide  
C/Species: Homo sapiens (man)  
C/Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 21-Jul-2000  
C/Accession: S07144; I52832; A60703  
R/Julien, J.P.; Grosfeld, F.; Yazdanbakhsh, K.; Flavell, D.; Meijer, D.; Mushynski, W.  
Biochim. Biophys. Acta 909, 10-20, 1987  
A/Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron organization  
A/Reference number: S07144; MUID:87214213; PMID:3034332  
A/Accession: S07144  
A/Molecule type: DNA  
A/Residues: 1-544 <JUL>  
A/Cross-references: EMBL:X05608; NID:g1495072; PIDN:CAA29097.1; PID:g1279504  
A/Note: The authors translated the codon ATG for residue 366 as Asn  
R/Pospelov, V.A.; Pospelova, T.V.; Julien, J.P.  
Cell Growth Differ. 5, 187-196, 1994  
A/Title: AP-1 and Krox-24 transcription factors activate the neurofilament light gene promoter  
A/Reference number: I52832; MUID:94235564; PMID:8180132  
A/Accession: I52832  
A/Status: preliminary; translated from GB/EMBL/DBDJ  
A/Molecule type: DNA  
A/Residues: 1-10 <POS>  
A/Cross-references: GB:S70309; NID:g547176; PIDN:AADI4057.1; PID:g4261757  
R/Nomata, Y.; Watanabe, T.; Wada, H.  
J. Biochem. 93, 825-831, 1983  
A/Title: Highly acidic proteins from human brain: purification and properties of Glu-50 protein  
A/Reference number: A60703; MUID:83265667; PMID:6135695  
A/Accession: A60703  
A/Molecule type: protein  
A/Residues: 469-472, 'D', 474 <NOM>  
A/Experimental source: Glu-50 brain peptide  
A/Note: this acidic protein is named for its greater than fifty per cent glutamic acid content  
C/Genetics:  
A/Gene: GDB:NEFL; NFL  
A/Cross-references: GDB:120227; OMIM:162280

A;Map position: 8p21-8p21  
A;Introns: 349/3; 391/2; 498/1  
C;Superfamily: cytoskeletal keratin  
C;Keywords: brain; coiled coil; intermediate filament  
F;469-544/Product: Glu-50 peptide #status predicted <E50>

Query Match 7.0%; Score 103; DB 2; Length 544;  
Best Local Similarity 24.2%; Pred. No. 6.9; Mismatches 87; Indels 58; Gaps 12;  
Matches 59; Conservative 30;

QY 100 FEVLEQA---TEIVKWPVTAERKRWL-----AESLRG--PALD--LMHIVQADNPS 146  
DB 285 FTVLTSAKNTDAVRAAKDEVSERULLKAKTLEIACRGWNALEKQLOELEDKQAD 344  
QY 147 ISVECLFAKQVFGSLSSRRTAQVRYLKYQE-----EGEK---- 183  
DB 345 ISAMQ--DTINKLENELRTTKSEMARYLKEYQOLLNVKMDIEIAAYRKLEGETRLS 402  
QY 184 -----VSAYVLRLETLRLKA-----VEKRAIPRRIADQVRLQVNMAGATLNM 226  
DB 403 FTSVGSITSGYSQSQVFGRSAYGGLQTSYLMSTRSPSYTYTSHVQEQTEVEETIA- 461  
QY 227 LMCRLRELKQDQPPPSFLELMKVIREE-EEEAEPENESIER-----PEERDGYGRWNH 279  
DB 462 --SKAEAKDEPPSGEAEERKDEEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAE 516  
QY 280 EGD 283  
DB 517 EGEE 520

RESULT 9  
T39530  
hypothetical protein SPBC1685.14c - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C;Accession: T39530  
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.  
submitted to the EMBL Data Library, July 1998  
A;Reference number: Z21861  
A;Accession: T39530  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-803 <WOO>  
A;Cross-references: EMBL:AL011154; PIDN:CAA20062.1; GSPDB:GN00067; SPDB:SPBC1685.14c  
A;Experimental source: strain 972h-; cosmid c1685  
C;Genetics:  
A;Map position: 2  
A;Gene: SPDB:SPBC1685.14c  
A;Introns: 8/2; 41/2

Query Match 7.0%; Score 103; DB 2; Length 803;  
Best Local Similarity 20.1%; Pred. No. 11; Mismatches 89; Indels 138; Gaps 15;  
Matches 67; Conservative 40;

QY 15 PNQDTEFLRLNLFLEKEGQTVSGMPFALQEGVSPATVPCISPELLAHLLGQMAHAPQ 74  
DB 169 PSKEVDSLENTVTKDLSE---DMLRKQEQDLDEIAK-----KYLQ--QEAEE 216  
QY 75 PLLPMRYKRLVFGSGAVPAPEE-----SFEVWLEQATEIV---- 111  
DB 217 PLVQQQ-----TSIVNPEKEEVTKTENIKSLEGELMTGISAEHLFDVAVEVFILQ 267  
QY 112 -----KEPVTVEAKRWLAELRGPALDLMHIVQAD--NPSISVEECLEAF 156  
DB 268 DPNVEASVFDLGDWNYFTISTEKTWLSQS-----VDADWNPVFSFE----- 310  
QY 157 KQVFGSLESRRTAQVRYLKYQEERKVSAYVLRLETLRLKAVEKRAIPRRIADQVRLQ 216  
DB 311 -----HLSFVWTFDANSNAFSLWLPDS-----QVRMEQ 340  
QY 217 VM-----AGATLQMLCRL-----RELKQDQPPPSFLELMKVIREE 255

DB 341 FOELLMRALWBSLNQORWLKIDDRQDYVMTTFHDEELED-SEDEEPARQOLLKKEEE 399  
QY 256 EEA-----SPENESIEEPERDDGYGRWNHE 280  
DB 400 EEEDEASDFEDSPADFDGEGADDLD-ESRWKE 432

RESULT 10  
T29091  
transitin - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T29091  
R;Cole, G.J.  
submitted to the EMBL Data Library, April 1996  
A;Reference number: Z20560  
A;Accession: T29091  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1840 <COL>  
A;Cross-references: EMBL:X80877; NID:g1279430; PID:e237634; PIDN:CAA56845.1

Query Match 7.0%; Score 103; DB 2; Length 1840;  
Best Local Similarity 25.3%; Pred. No. 32; Mismatches 87; Indels 82; Gaps 14;  
Matches 68; Conservative 32;

QY 20 EFLRLNLFLEKEGQTVS-----GMFRALQEGVSPATVPCISPELLAHLLGQA- 68  
DB 35 KFLSENEGIRAEIOSTKENPAGTHAGPGMRSSCGRSG-----CAAPRLHREVCGRAG 87  
QY 69 -----MAHAP---QPLPMRYKRLVFGSGAVPAPEESFEVWL-EGATEIVKEWPT 117  
DB 88 RDNLYBEVQHRVSRQCQEAEAEAKRQLSSSKLEEEERAAQIWLKERAVQLEKE---V 144  
QY 118 EAEKKRWLAESLRGALDLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRYLKY 177  
DB 145 EA-----LLEVHEEKAGLDQE--LASFSQ---SLEGRFCAVPVF---- 179  
QY 178 QBEQEKVSAYVLRLETLRLKAVEKRAIPRRIADQVRLQVNMAGATLQMLM----- 228  
DB 180 --QPVEDYSKRLSEIWRGAVETK-----AEVSQLERALQAKEN--LMQVAEDNQOS 230  
QY 229 -CRLRELKQDQPPPSFLELMKVIREEEE 256  
DB 231 QLQLRHLKE-----LVGLKVRKEMLEE 253

RESULT 11  
AB2304  
two-component hybrid sensor and regulator all3985 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C;Accession: AB2304  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AB2304  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-614 <KUR>  
A;Cross-references: GB:BA000019; PIDN:BA075684.1; PID:g17133120; GSPDB:GN00179  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: all3985

Query Match 7.0%; Score 102.5; DB 2; Length 614;  
Best Local Similarity 21.5%; Pred. No. 8.7; Mismatches 56; Conservative 34; Indels 81; Gaps 9;  
Matches 56;

QY 76 LLPMYRKLRFVSGSAV--PAPEESFEVWLEQATEI-----VKE 113



Db 356 RYEMLNRRMGSA--RQDDPSFEISDEILTHVAKS--VTASGRE-----LEG-AFNQLMF 404  
 QY 140 VQADNPSTSVRECLAEAFKQVFGSLRSRTAQVRYLKTQEGEKVSA---YVLRLETLIR 196  
 Db 405 RRSFEPNLSVDRVDLLSHLVGSGAKR---VRI-----EDIQIRIVARHYNVSRDELVS 456  
 QY 197 KAVEKRAIPRIA 209  
 Db 457 RRTVIVPRQIA 469

## RESULT 15

QPPGL  
 neurofilament triplet L protein - pig  
 N:Alternate names: 68K neurofilament protein  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 15-Nov-1984 #sequence revision 28-May-1986 #text\_change 10-Oct-1997  
 C:Accession: A91337; A90973; A34569; A02963  
 R:Geisler, N.; Plessmann, U.; Weber, K.  
 PDBS Lett. 182, 475-478, 1985  
 A:Title: The complete amino acid sequence of the major mammalian neurofilament protein  
 A:Reference number: A91337; MUID:85f54583; PMID:3920075  
 A:Accession: A91337  
 A:Molecule type: protein  
 A:Residues: 1-547 <GE1>  
 R:Geisler, N.; Kaufmann, E.; Fischer, S.; Plessmann, U.; Weber, K.  
 EMBO J. 2, 1295-1302, 1983  
 A:Title: Neurofilament architecture combines structural principles of intermediate filament  
 A:Reference number: A90973  
 A:Accession: A90973  
 A:Molecule type: protein  
 A:Residues: 1-82; 278-548 <GE2>  
 A:Note: Residue 322 is either lysine or arginine  
 R:Conda, Y.; Nishizawa, K.; Ando, S.; Kitamura, S.; Minoura, Y.; Nishi, Y.; Inagaki, M.  
 Biochem. Biophys. Res. Commun. 167, 1316-1325, 1990  
 A:Title: Involvement of protein kinase C in the regulation of assembly-disassembly of ne  
 A:Reference number: A34569; MUID:90211318; PMID:2108674  
 A:Accession: A34569  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 9-14; 23-29; 30-53 <GON>  
 C:Comment: Mammalian neurofilaments usually contain three polypeptides, L, M, and H (wit  
 ke all other intermediate filament proteins: a conserved alpha-helical region, whose hel  
 C:Comment: The amino-terminal headpiece is basic with a high content of hydroxyamino aci  
 al beta turns; domain b is acidic and rich in glutamic acid and lysine residues.  
 C:Comment: The extra mass and high charge density that distinguish the neurofilament pro  
 charged scaffolding structure suitable for interaction with other neuronal components  
 C:Comment: The boundaries of the domains between residues 70-92 and 399-402 are not yet  
 C:Comment: This protein was isolated from spinal cord.  
 C:Superfamily: cytoskeletal keratin  
 C:Keywords: coiled coil; intermediate filament  
 F:1-70/Domain: head <HED>  
 F:92-123/Domain: coil 1a, alpha-helical rod <R1A>  
 F:137-232/Domain: coil 1b, alpha-helical rod <R1B>  
 F:255-399/Domain: coil 2, alpha-helical rod <RD2>  
 F:402-548/Domain: tail <TAI>  
 F:402-442/Region: tail subdomain a  
 F:443-548/Region: tail subdomain b

Query Match 6.9%; Score 100.5; DB 1; Length 548;  
 Best Local Similarity 19.9%; Pred. No. 11;  
 Matches 56; Conservative 36; Mismatches 90; Indels 99; Gaps 11;  
 QY 80 RYKRLRVFGSAVPAPEESFEVWLEQATEIVKEWPTAEKKRWLAESLRGPALDLMHI 139  
 Db 263 QYEKL-----AAKMQNAEEWFKSRFTVLTE-----SAAKNTDAVRAKDEVSESRRL 310  
 QY 140 VQADNPSTSVRECLAEAFKQVFGSLRSRTAQVRYLKTQEGEKVSA---YVLRLETLIR 174  
 Db 311 LKA--KTLIEACXGWNLEAKQLQLEDKQADISAMQDTINKLENLRTTKSEMARYL 368  
 QY 175 KTYQE-----EGEK-----VSAYVLRLETLIRKA----- 198

Db 369 KEYQDLLNVKVALDIEIAAYRKLLEGEETRUSFTSVGSLTTGYSSQVFGRSAYGGLOT 428  
 QY 199 -----VEKRAIPRIADQVRLEQVWAGATLQMLMCRLELKDQGPSPSFLMLKVIRE -- 252  
 Db 429 SSYLMSTRSFPSYYTSHVQEQIEVEETIEA---AKAEAKDE--PPSEGEAEBSGKEKE 483  
 QY 253 -----EEEEASPFENBSIEEPERDGYGRWNHE 280  
 Db 484 EAAEAAEAEEGAQEEEAKESEEAKEBEGGEGGQGE 524

Search completed: September 21, 2004, 13:40:42  
 Job time : 24.3916 secs

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	1459	99.8	364	1	PNA2_HUMAN	Q9ul42 homo sapien
2	1428	97.7	364	1	PNA2_MACFA	Q9gm3 macaca fasc
3	1193.5	81.6	365	1	PNA2_MOUSE	Q9bhk0 mus musculus
4	634	43.4	353	1	PNA1_RAT	Q9vhz4 rattus norv
5	628	43.0	353	1	PNA1_HUMAN	Q8nd90 homo sapien
6	617	42.2	353	1	PNA1_MOUSE	Q9clc8 mus musculus
7	596.5	40.8	351	1	MOP1_HUMAN	Q9cby2 homo sapien
8	594	40.6	351	1	MOP1_MACFA	Q95ki4 macaca fasc
9	571.5	39.1	352	1	MOP1_MOUSE	Q9erh6 mus musculus
10	113	7.7	2442	1	CBP2_HUMAN	Q9bv73 homo sapien
11	112	7.7	554	1	NFL_HOVIN	P02548 bos taurus
12	110.5	7.6	544	1	NFL_XENLA	P35616 xenopus lae
13	110.5	7.6	1714	1	ITN1_MOUSE	Q9z0r4 mus musculus
14	108.5	7.4	1217	1	ITN1_RAT	Q9wve9 rattus norv
15	107	7.3	525	1	NAB2_MOUSE	Q6l127 mus musculus
16	107	7.3	1721	1	ITN1_HUMAN	Q15811 homo sapien
17	106	7.3	541	1	NFL_RAT	P19527 rattus norv
18	105.5	7.2	659	1	MX2_RAT	P18589 rattus norv
19	105.5	7.2	659	1	MX3_RAT	P18590 rattus norv
20	105	7.2	582	1	PESC_BRARE	P79741 brachydanio
21	105	7.2	759	1	CHLD_ARATH	Q9ej61 arabidopsis
22	104.5	7.1	525	1	NAB2_HUMAN	Q15742 homo sapien
23	103.5	7.1	548	1	NFL_FIG	P02547 sus scrofa
24	103	7.0	543	1	NFL_HUMAN	P07196 homo sapien
25	102.5	7.0	879	1	SYA_LISMO	Q8y722 listeria mo
26	101.5	6.9	487	1	DNAA_AGRIT5	Q8uih1 agrobacteri
27	101.5	6.9	1330	1	ITN1_VULVU	Q97961 vulpes vulp
28	100.5	6.9	558	1	YH32_AQUAE	O87622 aquifex aeo
29	100.5	6.9	879	1	SYA_ILISIN	Q92bk9 listeria in
30	99.5	6.8	542	1	NFL_MOUSE	P08551 mus musculus
31	99	6.8	399	1	RPSD_THEMEA	P77994 thermotoga
32	98.5	6.7	1962	1	MYSA_DROME	P05661 drosophila
33	98	6.7	1172	1	CNA2_MOUSE	Q80xc6 mus musculus

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF037365; AAD02098.1; -
DR EMBL; AF083114; AAF05625.1; -
DR EMBL; AF083115; AAF05626.1; -
DR EMBL; AF286487; AAG28165.1; -
DR EMBL; AB020690; BAA74906.1; -
DR GenBank; HGNC:9159; PMA2.
KW Antigen; Tumor antigen; Nuclear protein.
FT DOMAIN 333 338 POLY-GLU.
FT CONFLICT 128 129 GV -> AL (IN REF. 1; AAD02098).
FT CONFLICT 141 141 LV -> I (IN REF. 1; AAD02098).
FT CONFLICT 257 257 T -> P (IN REF. 1; AAF05625).
FT CONFLICT 278 278 R -> K (IN REF. 1; AAF05626).
SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96B3F0E93 CRC64;

Query Match 99.8%; Score 1459; DB 1; Length 364;
Best Local Similarity 99.6%; Pred. No. 1.1e-96;
Matches 282; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VQGGGVKVIPTNPQDTEFLERLNLFLEKGGQTVSGMFRALGQGVSPATVPCISPEL 60
Db 82 VQGGGVKVIPTNPQDTEFLERLNLFLEKGGQTVSGMFRALGQGVSPATVPCISPEL 141
QY 61 LAHLGQAMAHAPQPLPMRYKLVFSGSAPVAPPEESFEVWLEQATEIVKWPVTEAE 120
Db 142 LAHLGQAMAHAPQPLPMRYKLVFSGSAPVAPPEESFEVWLEQATEIVKWPVTEAE 201
QY 121 KKRWLAEISLRGPAIDLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRYLKYQEE 180
Db 202 KKRWLAEISLRGPAIDLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRYLKYQEE 261
QY 181 GKVSAYVLRLETLKXAVEKRAIPRIADQVRLQVMAQATLNQMLWCRRLKDKQGGP 240
Db 262 GKVSAYVLRLETLKXAVEKRAIPRIADQVRLQVMAQATLNQMLWCRRLKDKQGGP 321
QY 241 PSFLELMKVIREEESEAFSENEISIEPEERDGYGRWNHEGDD 283
Db 322 PSFLELMKVIREEESEAFSENEISIEPEERDGYGRWNHEGDD 364

RESULT 2
PMA2_MACFA
ID_PMA2_MACFA STANDARD; PRT; 364 AA.
AC Q9GMU3;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Paraneoplastic antigen Ma2 homolog.
GN PMA2.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Okada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the PMA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB047632; BAB12156.1; -
DR Nuclear protein.
FT DOMAIN 333 338 POLY-GLU.
SQ SEQUENCE 364 AA; 41950 MW; 0CF72210D7EC1524 CRC64;

Query Match 97.7%; Score 1428; DB 1; Length 364;
Best Local Similarity 97.2%; Pred. No. 1.8e-94;
Matches 275; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 VQGGGVKVIPTNPQDTEFLERLNLFLEKGGQTVSGMFRALGQGVSPATVPCISPEL 60
Db 82 VQGGGVKVIPTNPQDTEFLERLNLFLEKGGQTVSGMFRALGQGVSPATVPCISPEL 141
QY 61 LAHLGQAMAHAPQPLPMRYKLVFSGSAPVAPPEESFEVWLEQATEIVKWPVTEAE 120
Db 142 LAHLGQAMAHAPQPLPMRYKLVFSGSAPVAPPEESFEVWLEQATEIVKWPVTEAE 201
QY 121 KKRWLAEISLRGPAIDLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRYLKYQEE 180
Db 202 KKRWLAEISLRGPAIDLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRYLKYQEE 261
QY 181 GKVSAYVLRLETLKXAVEKRAIPRIADQVRLQVMAQATLNQMLWCRRLKDKQGGP 240
Db 262 GKVSAYVLRLETLKXAVEKRAIPRIADQVRLQVMAQATLNQMLWCRRLKDKQGGP 321
QY 241 PSFLELMKVIREEESEAFSENEISIEPEERDGYGRWNHEGDD 283
Db 322 PSFLELMKVIREEESEAFSENEISIEPEERDGYGRWNHEGDD 364

RESULT 3
PMA2_MOUSE
ID_PMA2_MOUSE STANDARD; PRT; 365 AA.
AC Q8BHK0;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Ma2 homolog.
GN PMA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain cortex;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogi A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bratt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Guatinchich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie I., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

```

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).  
 CC -!- SIMILARITY: Belongs to the PNMA family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AK043718; BAC11626.1; -;  
 DR EMBL; AK043910; BAC31700.1; -;  
 DR MGD; MGI:2444129; Pnma2.  
 KW Nuclear protein.  
 FT DOMAIN 333 338 POLY-GLU  
 SQ SEQUENCE 365 AA; 41201 MW; 0F90C940B9D843D9 CRC64;  
 -----  
 Query Match 81.6%; Score 1193.5; DB 1; Length 365;  
 Best Local Similarity 82.1%; Pred. No. 8.2e-78;  
 Matches 230; Conservative 21; Mismatches 28; Indels 1; Gaps 1;  
 -----  
 QY 1 VQGGVWVKVIFKTPNQDTEFLERLNLFLEKGGTQVSGMFRALGQGVSPATVPCISP 60  
 DB 1 VQGGVWVKVIFKTPNQDTEFLERLNLFLEKGGTQVSGMFRALGQGVSPATVPCISP 60  
 QY 82 VQGGVWVKVIFKTPNQDTEFLERLNLFLEKGGTQVSGMFRALGQGVSPATVPCISP 141  
 DB 82 VQGGVWVKVIFKTPNQDTEFLERLNLFLEKGGTQVSGMFRALGQGVSPATVPCISP 141  
 QY 61 LAHLGQMAHAPQPLT-PMRYKRLRVFSGSAPVAPPEESFEVWLEQATEIVKWPVTEAE 120  
 DB 142 LAHLTGQAMVHGQRPPLPVKCKMKRIFSGSTAAPEEPFVWLEQATEIAKEWPIPEAE 201  
 QY 121 KRWLAESLRGPALDLMHIVQADNPISIVBECLEAFKQVFGSLRRRTAQVRYLKYQEE 180  
 DB 202 KRWVAESLRGPALDLMHIVQADNPISIVBECLEAFKQVFGSLRRRTAQVRYLKYQEE 261  
 QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLQVWAGATINQMLWCRLEKQDQPP 240  
 DB 262 GEKISAVVLRLETLRLKAVEKRAIPRIADQVRLQVWAGATINQMLWCRLEKQDQPP 321  
 QY 241 PSFLELMKVIREEEEE-ASFENESIEEPEDRGYGRWNH 279  
 DB 322 PTFQLQMKVIREEEEEDAYFEQSRREPGRGSGCWNH 361  
 -----  
 RESULT 4  
 ID -PMAL RAT STANDARD; PRT; 353 AA.  
 AC Q8VHZ4;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Paraneoplastic antigen Mal homolog.  
 GN PNMA1 OR MAL.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 [1]  
 RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX MEDLINE=99158179; PubMed=10050892;  
 RA Dalmat J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,  
 RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frenn J.,  
 RA Posner J.B., Rosenfeld M.R.;  
 RT "Mal, a novel neuron- and testis-specific protein, is recognized by  
 RT the serum of patients with paraneoplastic neurological disorders.";  
 RT Brain 122:27-39 (1999).  
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).  
 CC -!- TISSUE SPECIFICITY: Testis and brain specific.  
 CC -!- SIMILARITY: Belongs to the PNMA family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF335505; AAL73196.1; -;  
 KW Nuclear protein.  
 FT DOMAIN 335 341 POLY-GLU  
 SQ SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEDA3 CRC64;  
 -----  
 Query Match 43.4%; Score 634; DB 1; Length 353;  
 Best Local Similarity 48.7%; Pred. No. 4.5e-38;  
 Matches 130; Conservative 51; Mismatches 74; Indels 12; Gaps 4;  
 -----  
 QY 3 GKGGVWVKVIFKTPNQDTEFLERLNLFLEKGGTQVSGMFRALGQGVSPATVPCISP 62  
 DB 83 GKGGLWVKVIFKTPNQDTEFLERLNLFLEKGGTQVSGMFRALGQGVSPATVPCISP 138  
 QY 63 HLLGQMAHAPQPLT-PMRYKRLRVFSGSAPVAPPEESFEVWLEQATEIVKWPVTEAE 121  
 DB 139 EMLNYLQNVIOPLVESIWKYKLTIFSGRDIPGCEETFDWLEHSNEVIEWQVSDIEK 198  
 QY 122 KRWLAESLRGPALDLMHIVQADNPISIVBECLEAFKQVFGSLRRRTAQVRYLKYQEE 181  
 DB 199 RRRLMESLRGPADVIRILKTNPAITTAECLEKALEQVFGSVSSRDQVRLNTYQNP 258  
 QY 182 EKVSAYVLRLETLRLKAVEKRAIPRIADQVRLQVWAGATINQMLWCRLEKQDQ 237  
 DB 259 EKLVSIVLRLEPLQKQVVDKGAIDKNVQARLQVIAAGNHSGALRQLML---TGATE 315  
 QY 238 GPPSFLELMKVIREEEEEASFENES 264  
 DB 316 GPAPNLQQLVQIREERAKEEEEA 342  
 -----  
 RESULT 5  
 ID -PMAL HUMAN STANDARD; PRT; 353 AA.  
 AC Q8ND90; O95144; Q8NG07;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Paraneoplastic antigen Mal (Neuron- and testis-specific protein 1)  
 DE (37 kDa neuronal protein).  
 GN PNMA1 OR MAL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
 RX MEDLINE=99158179; PubMed=10050892;  
 RA Dalmat J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,  
 RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frenn J.,  
 RA Posner J.B., Rosenfeld M.R.;  
 RT "Mal, a novel neuron- and testis-specific protein, is recognized by  
 RT the serum of patients with paraneoplastic neurological disorders.";  
 RT Brain 122:27-39 (1999).  
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).  
 CC -!- TISSUE SPECIFICITY: Testis and brain specific.  
 CC -!- SIMILARITY: Belongs to the PNMA family.

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [4]  
 RP SEQUENCE OF 162-353 FROM N.A.  
 RC TISSUE=Testis;  
 RA Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar. In tumor cells, it is  
 CC cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Testis and brain specific. In some patients  
 CC suffering from cancers, it is also specifically expressed by the  
 CC paraneoplastic tumor cells.  
 CC -!- MISCELLANEOUS: Antibodies against PNMAL are present in sera from  
 CC patients suffering of paraneoplastic neurological disorders.  
 CC -!- SIMILARITY: Belongs to the PNMA family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF037364; AAD13810.3; -  
 CC EMBL; AF320308; AAN05100.1; -  
 CC EMBL; BC039577; AAH39577.1; -  
 CC EMBL; AL834327; CAD38995.1; -  
 CC Genew; HGNC:9158; PNMA1.  
 CC MIM; 604010; -  
 CC GO; GO:0005737; C:cytoplasm; TAS.  
 CC GO; GO:0005730; C:nucleolus; TAS.  
 CC GO; GO:0007417; P:central nervous system development; TAS.  
 CC GO; GO:0007283; P:spermatogenesis; TAS.  
 CC Antigen; Tumor antigen; Nuclear protein.  
 CC DOMAIN 335 341 POLY-GLU.  
 CC SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;  
 CC  
 CC Query Match 43.08; Score 628; DB 1; Length 353;  
 CC Best Local Similarity 47.9%; Pred. No. 1.2e-37;  
 CC Matches 128; Conservative 52; Mismatches 75; Indels 12; Gaps 4;  
 CC  
 CC QY 3 GKGGVKVIKTPNQDTFLERLNLFLEKEGQTVSGMFRALGQGVGSPATVPCISPELLA 62  
 CC DB 83 GKGGVKVLKPPPTSDAEFLERHLFLAREGTVQDVARVLGFGNPPTP-----GPEMPA 138  
 CC  
 CC QY 63 HLLGQMAHAPOPLL-PMRYKRLRVFGSAVPAPEESFEVWLQATETIVKEHPVTEAK 121  
 CC DB 139 EMLNYILDNVQPLVESIWKYKRLTLFSGRDPGSGEETFPWLEHTNEVLEEWQSVDEK 198  
 CC  
 CC QY 122 KRWLAESIRGPDALMLHIVQADNPISVBECLFAKQVFGSLSRRTAQVRYLYKTYOEG 181  
 CC DB 199 RRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSVESDAQFLNTYQNPQ 258  
 CC  
 CC QY 182 EKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGA-----TLNQMLWCRRLRLKQ 237

Db 259 EKLSAYVLRLEPQLQVVEKGAIDKNVNVQARLEQVIAHNSGAIRQLWL---TGAGE 315  
 QY 238 GPPPSFLELMKVIREEEEESAPENES 264  
 Db 316 GPAPNLFOLLVQIREEEAKKEEEAEAA 342  
 RESULT 6  
 PNMA1 MOUSE  
 ID PNMA1 MOUSE STANDARD; PRT; 353 AA.  
 AC Q8CIC8; Q9CYP2;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Paraneoplastic antigen Mal homolog.  
 GN PNMA1 OR MAL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravagi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).  
 CC -!- SIMILARITY: Belongs to the PNMA family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AK017476; BAB30762.1; -  
 CC EMBL; AK028331; BAC25885.1; -  
 CC MGD; MGI:2180564; Pnmal.  
 CC Nuclear protein.  
 CC DOMAIN 336 341 POLY-GLU.  
 CC CONFLICT 136 136 T -> M (IN REF. 1; BAB30762).  
 CC SEQUENCE 353 AA; 39688 MW; 18CED3AC4E70939 CRC64;

Query Match 42.2%; Score 617; DB 1; Length 353;  
 Best Local Similarity 49.1%; Pred. No. 7.3e-37;  
 Matches 130; Conservative 47; Mismatches 72; Indels 16; Gaps 5;

QY 3 GKGWGVKVFPTNODTEFLERINLFLKEGQTVSGMFRALQGEVSPATVPCISPPELLA 62  
 DB 83 GKGGLMKVVFPTSDAVFLERLHLFLAREGTVQDVARVLGFQNPAPAP-----GPETPA 138

QY 63 HLLGQMAHAPQPLP--PMRYKLRVFGSAVPAPEEESFEVWLEQATEIVKEWPTAEK 121  
 DB 139 EMLYILNDVQLPVESTWYKLTUFGNDIPGGEETFDWLEHSENLIEBWQSDIEK 198

QY 122 KRWLAESLRGPALDLMHIVQADNPISVVEECLEAFKQVFGSLERRTAQVRVLYKTYQEG 181  
 DB 199 RRLMESLRGPAADVIRILKTNAITTAECLEKALEQVFGSVESRDQVRFELNTYQPG 258

QY 182 EKVSAYVLRLETLRKAVEKRAIPRIADQVLEQVMAGA-----TLNQLMWCRLRLKDKQ 237  
 DB 259 EKLSSVIRLEPLQKVVVDKGVIDKDNVQARLEQVIAGANHSGALRQLMLAGAE---E 315

QY 238 GPPPSFLELMKVIREE-----EEEE 258  
 DB 316 GPANFLQLLVQIREEAKKEEEA 340

## RESULT 7

MOPI HUMAN  
 ID MOPI HUMAN STANDARD; PRT; 351 AA.  
 AC Q96BY2; Q9H833; Q9HAS1;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Modulator of apoptosis 1 (MAP-1) (Paraneoplastic antigen MA4).  
 GN MAP1 OR PNMA4  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF LEU-120;  
 RP 120-LEU-ARG-127 AND 125-GLY-GLU-127.  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=21264738; PubMed=11060313;  
 RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C.,  
 RA Yu V.C.;  
 RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that  
 RT associates with Bax through its Bcl-2 homology domains.";  
 RL J. Biol. Chem. 276:2802-2807(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Retinoblastoma;  
 RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,  
 RA Wgatsuma M., Hosoari T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yanamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heileh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Ustin B., Toshynki S., Carninci P., Prange C.,  
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettman M., Madao A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Mediates caspase-dependent apoptosis.  
 CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX.  
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with higher levels in  
 CC heart and brain.  
 CC -!- DOMAIN: The BH3-like domain is required for association with BAX  
 CC and for mediating apoptosis. The three BH domains (BH1, BH2, and  
 CC BH3) of BAX are all required for mediating protein-protein  
 CC interaction.  
 CC -!- SIMILARITY: Belongs to the PNMA family.  
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a  
 CC frameshift in position 102.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; AF305550; AAC31786.1;  
 DR EMBL; AK024029; BAB14788.1; ALT\_SEQ.  
 DR EMBL; BC015044; AAH15044.1;  
 DR Genbank; HGNC:16658; MOAP1.  
 KW Apoptosis.  
 FT SITE 120 127 BH3-LIKE.  
 FT MUTAGEN 120 127 MISSING: ABROGATED INTERACTION WITH BAX,  
 FT MUTAGEN 120 127 RESULTING IN A NONAPOPTOTIC PROTEIN.  
 FT MUTAGEN 120 120 L->E: WEAKENED INTERACTION WITH BAX,  
 FT MUTAGEN 125 127 RESULTING IN A NONAPOPTOTIC PROTEIN.  
 FT CONFLICT 244 244 GHE->VLA: ABROGATED INTERACTION WITH BAX,  
 FT CONFLICT 258 258 T -> A (IN REF. 2).  
 FT CONFLICT 259 259 Y -> H (IN REF. 2).  
 FT CONFLICT 259 259 Q -> H (IN REF. 1).  
 SQ SEQUENCE 351 AA; 39512 MW; 5310142AC02B563C CRC64;

Query Match 40.8%; Score 596.5; DB 1; Length 351;  
 Best Local Similarity 48.3%; Pred. No. 2.1e-35;  
 Matches 128; Conservative 49; Mismatches 75; Indels 13; Gaps 5;

QY 1 VQKGGVKKVIFKTPNODTEFLERINLFLKEGQTVSGMFRALQGEVSPATVPCISP 60  
 DB 82 IPKGGGIVRVIFKPPDPDPTFLSRNLEFLAGEGTVGSLSPALCHENGSLDPEQGMIP 141

QY 61 LAHLLGQMAHAPQPLP--PMRYKLRVFGSAVPAPEEESFEVWLEQATEIVKEWPTAE 119  
 DB 142 WAPMLAQAL--EALQPALQCLKYLKRVFGSRESPEGESEFGRWFFHTTQMIKAWQPDV 200

QY 120 EKKELAESLRGPALDLMHIVQADNPISVVEECLEAFKQVFGSLERRTAQVRVLYKTYOE 179  
 DB 201 EKRRLESRGPALDVIRVLKNNPITVDECLQALEEVFGVTDNPRELQVRYLYTYQK 260

QY 180 EGEKVSAYVLRLETLRKAVEKRAIPRIADQVLEQVMAGA-----TLNQLMWCRLRLK 236  
 DB 261 DEEKLSAYVLRLEPLQLKLVQRGALRDVQARLDQVIAGAVHKTIRREL-----NLPE 315

QY 237 QGPPPSFLELMKVIREE-----EEEE 258  
 DB 316 DGPAPGLQLLVLIKDYEAEEEEEA 340

## RESULT 9



CC during interphase.  
 CC -!- SUBUNIT: Monomer and homodimer (Probable). Forms a complex in  
 CC vitro with both NEK2 kinase and the PP2C catalytic subunit of  
 CC protein phosphatase 1 (PP1).  
 CC -!- SUBCELLULAR LOCATION: Component of the core centrosome. In  
 CC interphase cells, it specifically associates with the proximal  
 CC ends of both mother and daughter centrioles. Associates with the  
 CC centrosome in interphase cells. In mitotic cells, it dissociates  
 CC from the mitotic spindle poles. At the end of cell division, it  
 CC reaccumulates at centrosomes.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=Q9BV73-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9BV73-2; Sequence=VSP\_007372;  
 CC Name=3;  
 CC IsoId=Q9BV73-3; Sequence=VSP\_007370, VSP\_007371;  
 CC Note=No experimental confirmation available;  
 CC TISSUE SPECIFICITY: Ubiquitously and weakly expressed.  
 CC -!- PTM: Differentially phosphorylated during cell cycle.  
 CC Phosphorylation may regulate association/dissociation from  
 CC centrosome. During M phase of mitosis, C-terminal part is  
 CC phosphorylated by NEK2, suggesting that it may trigger the  
 CC dissociation from the mitotic centrosome. It is dephosphorylated  
 CC in vitro by the PP1 phosphatase.  
 CC -!- DISEASE: Antibodies against CEP2 are present in sera from patients  
 CC with autoimmune diseases that developed autoantibodies against  
 CC centrosomal proteins.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF022655; AAC06349.1; -  
 CC EMBL; AF049105; AAC07988.1; -  
 CC EMBL; AL121586; CAB89415.1; -  
 CC EMBL; BC001433; AAH01433.1; -  
 CC Genew; HGNC:1859; CEP2  
 CC Cell cycle; Coiled coil; Phosphorylation; Alternative splicing;  
 CC Polymorphism.  
 CC DOMAIN 95 158 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 244 352 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 395 1172 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 1243 2227 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 2262 2376 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 246 250 POLY-LEU.  
 CC FT DOMAIN 464 2171 GLN/GLU-RICH.  
 CC FT VARSPLIC 313 313 V -> F (in isoform 3).  
 CC FT FTId=VSP\_007370.  
 CC FT Missing (in isoform 3).  
 CC FT FTId=VSP\_007371.  
 CC FT Missing (in isoform 2).  
 CC FT FTId=VSP\_007372.  
 CC FT Q -> H (in dbSNP:2296403).  
 CC FT FTId=VAR\_015649.  
 CC FT L -> I (IN REF. 1).  
 CC FT E -> A (IN REF. 1).  
 CC FT H -> L (IN REF. 1).  
 CC FT D -> E (IN REF. 1).  
 CC FT S -> D (IN REF. 2).  
 CC FT S -> I (IN REF. 1).  
 CC FT E -> A (IN REF. 1).  
 CC FT EGTK -> DEPO (IN REF. 1).  
 CC FT Q -> H (IN REF. 1).  
 CC FT CONFLICT 1153 1153  
 CC FT CONFLICT 1246 1246  
 CC FT CONFLICT 1513 1513  
 CC FT CONFLICT 2082 2082  
 CC FT CONFLICT 2345 2345

CC SQ SEQUENCE 2442 AA; 281135 MM; BC2B8A36E07B8272 CRC64;  
 CC Query Match 7.7%; Score 113; DB 1; Length 2442;  
 CC Best Local Similarity 22.7%; Pred. No. 4.9;  
 CC Matches 75; Conservative 46; Mismatches 91; Indels 118; Gaps 16;  
 CC  
 CC QY 22 LERLMLFLEKGGQTVSGMFRALGQEGVSPATVPCLSPPELLAHLLGQMAHAPQPLLPWRY 81  
 CC 1458 LELLSLDLKKNQVD-----LQRFQIQELE-KCRS--VLEH-----LPMVA 1496  
 CC  
 CC QY 82 RKLRFVSGSAPPAPEESFEVLEQATIVK-----EWPVTEAEKKRWLAESLRGP 132  
 CC 1497 QE-----REQKLTVOREQIRELEKDRFQNRVLEHQHLELEKQDMIESQRQ 1544  
 CC  
 CC QY 133 ALDL-MHIVQADNPISVE-----ECLE-AFKQVFGSLESRTA----- 169  
 CC 1545 VQDLKKQLVTLCLALEENHHKMECCQKLIKELGGQRETQRTVALTHLTLDLEERSQBL 1604  
 CC  
 CC QY 170 -----QVRYLKY-----QEGEKYSAYVIRLETILRKA-----VEKR----- 202  
 CC 1605 QAQSSQIHDLSEHSTVLARELQERDQVKSQREIEEQRQKEHLTQDLERRDQELMLQK 1664  
 CC  
 CC QY 203 -----ALPRIAQVRLQVMAG---ATLNQMLWCRLRELKDQGPPTS--- 242  
 CC 1665 ERIVQLEDORTQTKLEEDLEQIKLSLEERGRELTTQQL---MQERAEEGKGPSXAK 1720  
 CC  
 CC QY 243 --FLELMKVIREEESEASFENESIEEPPEE 270  
 CC 1721 RGSLEHMKLILRDKEKEVCEQCEHIELQE 1750  
 CC DB  
 CC  
 CC RESULT 11  
 CC ID\_NFL\_BOVIN STANDARD; PRT; 554 AA.  
 CC AC P02548; P79127;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 CC DE Neurofilament triplet L protein (68 kDa neurofilament protein)  
 CC DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich  
 CC protein).  
 CC GN NEFL.  
 CC OS Bos taurus (Bovine).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 CC OC Bovidae; Bovinae; Bos.  
 CC OX NCBI\_TaxID=9913;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=Holstein; TISSUE=Brain;  
 CC RA Hill W.D., Zhang L., Balin B.J., Sprinkle T.J.;  
 CC RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC RN [2]  
 CC RP SEQUENCE OF 468-554.  
 CC RX MEDLINE=85154567; PubMed=3884373;  
 CC RA Isobe T., Okuyama T.;  
 CC RT "Brain micro glutamic acid-rich protein is the C-terminal endpiece of  
 CC the neurofilament 68-kDa protein as determined by the primary  
 CC sequence.";  
 CC RL FES Lett. 182:389-392(1985).  
 CC -!- FUNCTION: Neurofilaments usually contain three intermediate  
 CC filament proteins: L, M, and H which are involved in the  
 CC maintenance of neuronal caliber.  
 CC -!- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH  
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT  
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM  
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH  
 CC OTHER NEURONAL COMPONENTS OR IONS.  
 CC -!- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE  
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL  
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM  
 CC FILAMENTS.  
 CC -!- SIMILARITY: Belongs to the intermediate filament family.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL: U83919; AAB41543.1; -  
 InterPro: IPR006821; Filament\_head.  
 InterPro: IPR001664; IF.  
 Pfam: PF00038; filament; 1.  
 Pfam: PF04732; filament head; 1.  
 PROSITE: PS00226; IF; 1.  
 Intermediate filament; Coiled coil; Neurone.  
 INIT MET 0 BY SIMILARITY.  
 FT DOMAIN 1 92 HEAD (BY SIMILARITY).  
 FT DOMAIN 93 396 ROD (BY SIMILARITY).  
 FT DOMAIN 397 554 TAIL (BY SIMILARITY).  
 FT DOMAIN 93 124 COIL 1A.  
 FT DOMAIN 125 137 LINKER 1.  
 FT DOMAIN 138 233 COIL 1B.  
 FT DOMAIN 234 252 LINKER 12.  
 FT DOMAIN 253 271 COIL 2A.  
 FT DOMAIN 272 280 LINKER 2.  
 FT DOMAIN 281 396 COIL 2B.  
 FT DOMAIN 397 443 TAIL, SUBDOMAIN A.  
 FT DOMAIN 444 554 TAIL, SUBDOMAIN B (ACIDIC).  
 FT CONFLICT 494 500 MISSING (IN REF. 2).  
 FT CONFLICT 509 509 A -> AEA (IN REF. 2).  
 SQ SEQUENCE 554 AA; 62514 MW; D772B81CA2C31C1A CRC64;

Query Match 7.7%; Score 112; DB 1; Length 554;  
 Best Local Similarity 24.2%; Pred. No. 0.97;  
 Matches 55; Conservative 30; Mismatches 82; Indels 60; Gaps 10;  
 Qy 100 FEVLEQA---TEIVKWPVTEAEKRWL-----AESLRG--PALD--LMHIVQADNPS 146  
 Db 284 FTVLTESAANKTDVRAAKDVESSRLAKAKTLEIEACRGWNALEKQLOLEDKQAD 343  
 Qy 147 ISVEECLEAFKQVFGSLERRTAQVRYLKYOE-----EGEK----- 183  
 Db 344 ISAMQ---DTINKLENELRTTKSEMARYLKEYQDLNVKQALDIEIAAYRKLEGEETRLS 401  
 Qy 184 -----VSAYVLRLETLRKA-----VKRAIPRIADQVRLQVWAGATLNQM 226  
 Db 402 FTSVGSUTTGTSQVFGSAYGGLQTSYLMARSFPSPSYTSHVQEQIEVEETIEA- 460  
 Qy 227 LWCRLRELKQGPFPFLELMKLVIREEEERASPENESIERPERDQ 273  
 Db 461 --AKAEAKDE--PPEGEAEKEKEKEAEAEAEAEAEAEAEAE 503

RESULT 12  
 NFL XENLA  
 ID NFL XENLA STANDARD; PRT; 544 AA.  
 AC P35616;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-L).  
 DE L.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RT TISSUE=Brain;  
 RX MEDLINE=92356194; PubMed=1494944;  
 RA Charnas L.R., Szaro B.G., Gainer H.;

RT RT Identification and developmental expression of a novel low molecular weight neuronal intermediate filament protein expressed in Xenopus laevis.";  
 RL J. Neurosci. 12:3010-3024(1992).  
 CC -!- FUNCTION: Neurofilaments usually contain three intermediate filament proteins: L, M, and H which are involved in the maintenance of neuronal caliber.  
 CC -!- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM FILAMENTS.  
 CC -!- SIMILARITY: Belongs to the intermediate filament family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 DR EMBL: M86654; AAB3018.1; -  
 DR InterPro: IPR006821; Filament\_head.  
 DR InterPro: IPR001664; IF.  
 DR Pfam: PF00038; filament; 1.  
 DR Pfam: PF04732; filament\_head; 1.  
 DR PROSITE: PS00226; IF; 1.  
 KW Intermediate filament; Coiled coil; Neurone.  
 FT DOMAIN 1 87 HEAD.  
 FT DOMAIN 88 390 ROD.  
 FT DOMAIN 391 544 TAIL.  
 FT DOMAIN 88 119 COIL 1A.  
 FT DOMAIN 120 132 LINKER 1.  
 FT DOMAIN 133 228 COIL 1B.  
 FT DOMAIN 229 246 LINKER 12.  
 FT DOMAIN 247 265 COIL 2A.  
 FT DOMAIN 266 274 LINKER 2.  
 FT DOMAIN 275 390 COIL 2B.  
 FT DOMAIN 391 435 TAIL, SUBDOMAIN A.  
 FT DOMAIN 436 544 TAIL, SUBDOMAIN B (ACIDIC).  
 FT DOMAIN 441 538 GLU-RICH.  
 FT DOMAIN 464 469 POLY-GLU.  
 SQ SEQUENCE 544 AA; 61861 MW; 76D911B896E97201 CRC64;

Query Match 7.6%; Score 110.5; DB 1; Length 544;  
 Best Local Similarity 22.3%; Pred. No. 1.2;  
 Matches 59; Conservative 30; Mismatches 83; Indels 93; Gaps 10;  
 Qy 80 RYKRLRVFGSAVPAPPEESFEVWLE-----QATEIVKWPVTEAEKRWLA-- 126  
 Db 258 QYEKL-----AAKQWQAEEWFKSRFTVLTOSAAARNTDAVRAAKDENSESRRLMSAK 309  
 Qy 127 ----ESLRGPALDLMHIVQ--ADNPSISVEECLEAFKQVFGSLERRTAQVRYLKYOE- 179  
 Db 310 GLEIEACRGVNEALQRQIQELEDKQSGEIAQMADINKLEELRNTKSEMARYLKEYQDL 369  
 Qy 180 -----EGEKV-----SAYVLRLETLRKA 198  
 Db 370 LNVKQALDIEIAAYRKLEGEETRLSFGVCAITSGYTSAPVFGRSAYSLQSSYV--- 426  
 Qy 199 VKRAIPRIADQVRLQVWAGATLNQMLWCLRELKQGPFPFLELMKLVIREEEEREA 258  
 Db 427 -TSRAFTYSSHVQEQEQLDIEETIS---SRAEAKAEAP-----EEEEEA 471  
 Qy 259 SFENESIEEPEDRDGYGRWNHGGD 283  
 Db 472 EEEGEGEAEAE-----GEEGEE 490

Query Match 7.6%; Score 110.5; DB 1; Length 544;  
 Best Local Similarity 22.3%; Pred. No. 1.2;  
 Matches 59; Conservative 30; Mismatches 83; Indels 93; Gaps 10;

Qy 80 RYKRLRVFGSAVPAPPEESFEVWLE-----QATEIVKWPVTEAEKRWLA-- 126  
 Db 258 QYEKL-----AAKQWQAEEWFKSRFTVLTOSAAARNTDAVRAAKDENSESRRLMSAK 309  
 Qy 127 ----ESLRGPALDLMHIVQ--ADNPSISVEECLEAFKQVFGSLERRTAQVRYLKYOE- 179  
 Db 310 GLEIEACRGVNEALQRQIQELEDKQSGEIAQMADINKLEELRNTKSEMARYLKEYQDL 369  
 Qy 180 -----EGEKV-----SAYVLRLETLRKA 198  
 Db 370 LNVKQALDIEIAAYRKLEGEETRLSFGVCAITSGYTSAPVFGRSAYSLQSSYV--- 426  
 Qy 199 VKRAIPRIADQVRLQVWAGATLNQMLWCLRELKQGPFPFLELMKLVIREEEEREA 258  
 Db 427 -TSRAFTYSSHVQEQEQLDIEETIS---SRAEAKAEAP-----EEEEEA 471  
 Qy 259 SFENESIEEPEDRDGYGRWNHGGD 283  
 Db 472 EEEGEGEAEAE-----GEEGEE 490

RESULT 13  
 ITN1 MOUSE  
 ID ITN1 MOUSE STANDARD; PRT; 1714 AA.  
 AC Q9Z0R4; Q9R143;

DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Intersectin 1 (EH and SH3 domains protein 1).  
 GN ITSN1 OR ITSN OR ESEL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=99164083; PubMed=10064583;  
 RA Sengar A.S., Wang W., Bishay J., Cohen S., Egan S.E.;  
 RT "The EH and SH3 domain ESE proteins regulate endocytosis by linking to  
 RT dynamin and Eps15.";  
 RL EMBO J. 18:11559-1171 (1999).  
 [2]  
 RP SEQUENCE OF 966-1714 FROM N.A. (ISOFORM 2), AND SEQUENCE OF 545-599  
 RP FROM N.A. (ISOFORMS 1 AND 2).  
 RC STRAIN=129/Ola; TISSUE=Spleen;  
 RA Skripkina I.Y., Tsyba L.O., Anoprienko O.V., Slavov D., Tassone F.,  
 RA Rymditch A.V., Gardiner K.;  
 RT "Mouse homologues of human chromosome 21 genes.";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Adapter protein that may provide indirect link between  
 CC the endocytic membrane traffic and the actin assembly machinery.  
 CC May regulate the formation of clathrin-coated vesicles.  
 CC -!- SUBUNIT: Interacts with dynamin, SNAP-25 and SNAP-23. Clusters  
 CC several dynamin in a manner that is regulated by alternative  
 CC splicing. Also binds clathrin-associated proteins and other  
 CC components of the endocytic machinery, such as N-WASP, Eps15 and  
 CC Stoinin 2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated protein.  
 CC Enriched in synaposomes (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1; Synonyms=Esel;  
 CC IsoId=Q9Z0R4-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9Z0R4-2; Sequence=VSP\_004296;  
 CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed at high levels in  
 CC brain, heart and skeletal muscle.  
 CC -!- DOMAIN: SH3-3, SH3-4 and SH3-5, but not SH3-1 and SH3-2 domains,  
 CC bind to dynamin (By similarity).  
 CC -!- DOMAIN: The KLERQ domain binds to SNAP-25 and SNAP-23 (By  
 CC similarity).  
 CC -!- MISCELLANEOUS: Overexpression results in the inhibition of the  
 CC transferrin uptake and the blockage of the clathrin-mediated  
 CC endocytosis.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.  
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.  
 CC -!- SIMILARITY: Contains 2 EH domains.  
 CC -!- SIMILARITY: Contains 1 PH domain.  
 CC -!- SIMILARITY: Contains 5 SH3 domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF132481; AAD19749.1; -;  
 CC EMBL; AF132478; AAD19746.1; -;  
 CC EMBL; AF169621; AAD48848.1; -;  
 CC EMBL; AF356517; AAK40228.1; -;  
 CC HSSP; P29354; IGFC  
 CC MGD; MGI:1338069; Itsn.  
 DR GO; GO:0030027; C:lamellipodium; IDA.  
 DR GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0007264; P:small GTPase mediated signal transduction; IDA.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR008973; C2 CalB.  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR000261; Eps15\_homology.  
 DR InterPro; IPR001849; PH.  
 DR InterPro; IPR000219; RhoGEF.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF00168; C2\_1.  
 DR Pfam; PF00036; efhand; 3.  
 DR Pfam; PF00169; PH; 1.  
 DR Pfam; PF00621; RhoGEF; 1.  
 DR Pfam; PF00018; SH3; 5.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 5.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00054; Efh; 2.  
 DR SMART; SM00027; EH; 2.  
 DR SMART; SM00233; PH; 1.  
 DR SMART; SM00325; RhoGEF; 1.  
 DR SMART; SM00326; SH3; 5.  
 DR PROSITE; PS50004; C2 DOMAIN 2; 1.  
 DR PROSITE; PS50010; DH\_2; 1.  
 DR PROSITE; PS00018; EF\_HAND; 2.  
 DR PROSITE; PS00031; EH; 2.  
 DR PROSITE; PS50003; PH DOMAIN; 1.  
 DR PROSITE; PS50002; SH3; 5.  
 KW Endocytosis; SH3 domain; Repeat; Coiled coil; Calcium-binding;  
 FT Alternative splicing.  
 FT DOMAIN 21 109  
 FT CA\_BIND 66 78  
 FT DOMAIN 221 310  
 FT CA\_BIND 267 279  
 FT CA\_BIND 326 702  
 FT DOMAIN 352 662  
 FT DOMAIN 738 799  
 FT DOMAIN 906 964  
 FT DOMAIN 995 1053  
 FT DOMAIN 1067 1131  
 FT DOMAIN 1148 1207  
 FT DOMAIN 1230 1416  
 FT DOMAIN 1455 1564  
 FT DOMAIN 1576 1672  
 FT DOMAIN 321 324  
 FT VARSPLIC 1214 1714  
 FT SEQUENCE 1714 AA; 194284 MW; 4D7AF298397860A7 CRC64;  
 Query Match 7.6%; Score 110.5; DB 1; Length 1714;  
 Best Local Similarity 26.0%; Pred. No. 4.8;  
 Matches 60; Conservative 23; Mismatches 61; Indels 87; Gaps 9;  
 QY 60 LLAHLGQAWAHAPQ-----LLPMRYKLRVFGSAVPA-----PEESFEVWLE 105  
 DB 281 LAWHLIDVAMSGQPLDPVLPPEYIPFSFRVSGSGMSVSSSVQRIPEESSED--E 338  
 QY 106 QATEIVKPEVTEAEKKRWLAESLRGALDMLHIVQADNPSISVEECLEAFKQVFGSLES 165  
 DB 339 QQPE--KLPVTFEDKKR-----ENFERGSEVLEK 366  
 QY 166 RRTAQVRYKYTKTQBEKEKVSAYVRLTLLRKAVEKRAIPRRITADQVRLQVWAGATLQ 225  
 DB 367 RQALLLEQQRKEQ-----RLAQLERAFQERK-----ERERQEQ----- 400  
 QY 226 MLWCRLRELKDGQPPPSFLELMKVIREEEERASFENESIEEPEEDDGYGR 276  
 DB 401 -----EAKRQ-----LELEKQLEKQRELEKQREERKEEREAACKR 438  
 RESULT 14  
 ITN1 RAT  
 ID ITN1 RAT STANDARD; PRT; 1217 AA.  
 AC Q9WVE9; Q9WVE1;

DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Intersectin 1 (EH domain and SH3 domain regulator of endocytosis 1).  
 GN ITSN1 OR ITSN OR EHSH1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=99303609; PubMed=10373452;  
 RA Okamoto M., Schoch S., Suedhof T.C.;  
 RT "EHSH1/Intersectin, a protein that contains EH and SH3 domains and  
 RT binds to dynamin and SNAP-25. A protein connection between exocytosis  
 RT and endocytosis?";  
 RL J. Biol. Chem. 274:18446-18454 (1999).  
 CC -!- FUNCTION: Adapter protein that may provide indirect link between  
 CC the endocytic membrane traffic and the actin assembly machinery.  
 CC May regulate the formation of clathrin-coated vesicles.  
 CC -!- SUBUNIT: Interacts with dynamin, SNAP-25 and SNAP-23. Clusters  
 CC several dynamin in a manner that is regulated by alternative  
 CC splicing. Also binds clathrin-associated proteins and other  
 CC components of the endocytic machinery, such as N-WASP, Eps15 and  
 CC stonin 2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated protein.  
 CC -!- Enriched in synapses.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1;  
 CC IsoId=Q9WVE9-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9WVE9-2; Sequence=VSP\_004297;  
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain.  
 CC -!- DOMAIN: SH3-3, SH3-4 and SH3-5, but not SH3-1 and SH3-2 domains,  
 CC bind to dynamin.  
 CC -!- DOMAIN: The KLRQ domain binds to SNAP-25 and SNAP-23.  
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.  
 CC -!- SIMILARITY: Contains 2 EH domains.  
 CC -!- SIMILARITY: Contains 5 SH3 domains.  
 CC -!- SIMILARITY: Contains 5 SH3 domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF127798; AAD30271.1; -;  
 DR EMBL; AF132672; AAD31026.1; -;  
 DR HSP; P29354; LGFC.  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR000261; Eps15\_homology.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF00036; efhand; 3.  
 DR Pfam; PF00018; SH3; 5.  
 DR PRINTS; PD00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 5.  
 DR SMART; SM00054; EFh; 2.  
 DR SMART; SM00027; EH; 2.  
 DR SMART; SM00326; SH3; 5.  
 DR PROSITE; PS00018; EF\_HAND; 2.  
 DR PROSITE; PS50031; EH; 2.  
 DR PROSITE; PS50002; SH3; 5.  
 DR Endocytosis; SH3 domain; Repeat; Coiled coil; Calcium-binding;  
 KW Alternative splicing  
 FT DOMAIN 21 109 EH 1.  
 FT CA\_BIND 66 78 EF\_HAND 1 (POTENTIAL).  
 FT DOMAIN 221 310 EH 2.  
 FT CA\_BIND 267 279 EF\_HAND 2 (POTENTIAL).

FT DOMAIN 326 702 LYS/LEU/GLU/ARG/GLN-RICH (KLRQ).  
 FT DOMAIN 350 670 COILED COIL (POTENTIAL).  
 FT DOMAIN 737 803 SH3 1.  
 FT DOMAIN 910 968 SH3 2.  
 FT DOMAIN 999 1057 SH3 3.  
 FT DOMAIN 1071 1135 SH3 4.  
 FT DOMAIN 1152 1211 SH3 5.  
 FT DOMAIN 321 324 POLY-SER.  
 FT VARSPLIC 1003 1073 Missing (in isoform 2).  
 FT SEQUENCE 1217 AA; 137154 MW; 6C13238AE5A5B34B CRC64;  
 SQ SEQUENCE 1217 AA; 137154 MW; 6C13238AE5A5B34B CRC64;  
 Query Match 7.4%; Score 108.5; DB 1; Length 1217;  
 Best Local Similarity 24.7%; Pred. No. 4.4;  
 Matches 57; Conservative 25; Mismatches 62; Indels 87; Gaps 8;  
 QY 60 LIAHLGLQMAHAPOQ-----LLPMRYRKLKRVFSGSAVPA-----PEESFEVWLE 105  
 Db 281 LAMHLIDVAMSGQPLPPVLPPEYIPSPFRVRSGSMVSSSSADQRLPEPS-----SE 336  
 QY 106 QATEIVKWPVTEAEKKWLAESLRGPDALDLMHIVQADNPISVBECLAEAFKQVFGSLES 165  
 Db 337 DEQVEKKLPVTFEDKKR-----ENPERGNLELEK 366  
 QY 166 RTAQVRYLKYOEKGVSAVYVLETLRLKAVEKRAIPRIADQVLEQVMAGATLNQ 225  
 Db 367 RQALLEQQRKEQ-----RLAQLEAERK-----ERERQEQ 400  
 QY 226 MLWCRLRELKDGPPPSFLMKVIRBEESAEAGFENESIEPERDGYR 276  
 Db 401 -----ERKEQ-----LELEKQLEKQRELEKQREERKEIRERAAKR 438  
 RESULT 15  
 NAB2 MOUSE  
 ID NAB2 MOUSE STANDARD; PRT; 525 AA.  
 AC Q61127;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE NGFI-A binding protein 2 (EGR-1 binding protein 2).  
 GN NAB2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC TISSUE=Brain;  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RX MEDLINE=96251303; PubMed=8668170;  
 RA Svaren J., Severson B.R., Apel E.D., Zimonjic D.B., Popescu N.C.,  
 RA Milbrandt J.;  
 RT "NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by  
 RT proliferative and differentiative stimuli";  
 RL Mol. Cell. Biol. 16:3545-3553 (1996).  
 RN [2]  
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=97271553; PubMed=9126479;  
 RA Svaren J., Apel E.D., Simburger K.S., Jenkins N.A., Gilbert D.J.,  
 RA Copeland N.A., Milbrandt J.;  
 RT "The Nab2 and Stat6 genes share a common transcription termination  
 RT region";  
 RL Genomics 41:33-39 (1997).  
 CC -!- FUNCTION: Acts as a transcriptional repressor for zinc finger  
 CC transcription factors EGR1 and EGR2. Isoform 2 lacks repression  
 CC ability.  
 CC -!- SUBUNIT: Homomultimeris may associate with EGR1 bound to DNA (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR. ISOFORM 2 IS NOT LOCALIZED TO THE  
 CC NUCLEUS.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;

```
CC      Isoid=Q61127-1; Sequence=Displayed;
CC      Name=2;
CC      Isoid=Q61127-2; Sequence=VSP_003388, VSP_003389;
CC      TISSUE SPECIFICITY: Highly expressed in brain and thymus, and at
CC      lower levels in spleen, kidney, heart and testis. Isoform 1 is
CC      predominantly expressed in testis, whereas isoform 3 is more
CC      abundant in thymus.
CC      INDUCTION: By serum stimulation.
CC      DOMAIN: The NAB conserved domain 1 (NCD1) interacts with EGRI
CC      inhibitory domain and mediates multimerization.
CC      DOMAIN: The NAB conserved domain 2 (NCD2) is necessary for
CC      transcriptional repression.
CC      SIMILARITY: BELONGS TO THE NAB FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U47543; AAC52650.1; -.
CC      MGD; MGI:107563; Nab2.
CC      InterPro; IPR006989; Nab centr.
CC      InterPro; IPR006988; Nab_N.
CC      Pfam; PF04904; NCD1; 1.
CC      Pfam; PF04905; NCD2; 1.
CC      Transcription regulation; Repressor; Alternative splicing.
CC      FT DOMAIN 35 113 NCD1.
CC      FT DOMAIN 267 356 NCD2.
CC      FT DOMAIN 353 384
CC      FT VARSPLIC 320 322 LTI -> ASL (in isoform 2).
CC      FT VARSPLIC 323 525 /FTId=VSP_003388.
CC      FT VARSPLIC 323 525 Missing (in isoform 2).
CC      FT /FTId=VSP_003389.
CC      SQ SEQUENCE 525 AA; 56653 MW; FED428E94A8BD804 CRC64;
CC
Query Match 7.3%; Score 107; DB 1; Length 525;
Best Local Similarity 22.6%; Pred. No. 2.1;
Matches 47; Conservative 39; Mismatches 84; Indels 38; Gaps 6;
QY 86 VFSGSAVPAP---EESFEVWLEQATEIVKEWPTVEAKKRWLAESLRGPALDLMHVOA 142
Db 225 VAAGGAGGCDRLPEPWRVWVESVERIFPSFPRGTGETASLLKLNKLARSVGHIFEM 284
QY 143 DNPISIVVECLEAFKQVFGSLESRRTAQVRYLKYOESEKVS-----AVLRL 191
Db 285 DDHDAQKEEBIRKYSVIYGRDLSK-----REGKQLSLHELTINEAAAAQFCMRD 333
QY 192 ETLLKAVKRAIPRIADQVRLEQVMAGNTLN--QMLWCLRELKD-----QGP 239
Db 334 NTLRLRVVLSLRQVARESTYLSLKGSLRLSHELGGPPLKKLQEVGEQSHNIQPP 393
QY 240 PPSFLELMKVIREE-EEESASFENESIE 266
Db 394 PPGPESYAPPYPSLEEDSASLSGESLD 421
```

Search completed: September 21, 2004, 13:32:39  
Job time : 9.71499 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 65.5774 Seconds  
(without alignments)  
1361.621 Million cell updates/sec

Title: US-10-037-860-11

Perfect score: 1462

Sequence: 1 VQGGGVWVKVIFKTPNQDTE.....SIEPEERDGYGRWNHEGDD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1462	100.0	283	4	Q9UL42	Q9ul42 homo sapien
2	1459	99.8	364	4	Q94959	Q94959 homo sapien
3	1428	97.7	364	6	Q9GMJ3	Q9gmj3 macaca fasc
4	1193.5	81.6	365	11	Q8BHK0	Q8bhk0 mus musculus
5	755	51.6	149	4	Q9UL43	Q9ul43 homo sapien
6	634	43.4	353	11	Q8VH24	Q8vh24 rattus norv
7	628	43.0	353	4	Q95144	Q95144 homo sapien
8	628	43.0	353	4	Q8NG07	Q8ng07 homo sapien
9	622	42.5	466	11	Q8JZW8	Q8jzw8 mus musculus
10	620	42.4	353	11	Q9CYP2	Q9cyp2 mus musculus
11	618.5	42.3	455	4	Q9H0A4	Q9h0a4 homo sapien
12	618.5	42.3	463	4	Q9UL41	Q9ul41 homo sapien
13	617	42.2	353	11	Q8CIC8	Q8cic8 mus musculus
14	597	40.8	448	4	Q8NET3	Q8net3 homo sapien
15	597	40.8	452	4	Q96PV4	Q96pv4 homo sapien
16	596.5	40.8	351	4	Q96BY2	Q96by2 homo sapien

17	594	40.6	351	6	Q95KI4	Q95ki4 macaca fasc
18	591.5	40.5	351	4	Q9HAS1	Q9has1 homo sapien
19	576	39.4	194	4	Q95145	Q95145 homo sapien
20	571.5	39.1	352	11	Q9ERH6	Q9erh6 mus musculus
21	480.5	32.9	237	4	Q9H833	Q9h833 homo sapien
22	475.5	32.5	399	4	Q96A40	Q96a40 homo sapien
23	453.5	31.0	192	4	Q8ND90	Q8nd90 homo sapien
24	403.5	27.6	402	11	Q9CZA5	Q9czas mus musculus
25	394	26.9	403	4	Q8TE36	Q8te36 homo sapien
26	391.5	26.8	393	11	Q9DB17	Q9db17 mus musculus
27	391.5	26.8	393	11	Q8VD24	Q8vd24 mus musculus
28	380.5	26.0	402	4	Q8N1C1	Q8n1c1 homo sapien
29	147	10.1	430	11	Q80VM8	Q80vm8 mus musculus
30	137.5	9.4	246	11	Q8C533	Q8c533 mus musculus
31	135	9.2	378	4	Q8N3H4	Q8n3h4 homo sapien
32	135	9.2	435	4	Q8GV59	Q8gv59 homo sapien
33	117	8.0	538	4	Q9H0W5	Q9h0w5 homo sapien
34	117	8.0	538	4	Q8TB26	Q8tb26 homo sapien
35	116.5	8.0	1220	13	Q8JFT5	Q8jft5 brachydanio
36	116.5	8.0	1721	13	Q8JFT4	Q8jft4 brachydanio
37	111	7.6	1150	17	Q8U256	Q8u256 pyrococcus
38	110	7.5	281	16	Q7VNB6	Q7vnb6 haemophilus
39	110	7.5	386	4	Q8TCR7	Q8tcr7 homo sapien
40	110	7.5	956	13	Q7ZW40	Q7zw40 brachydanio
41	109.5	7.5	1202	13	Q7ZYV9	Q7zyv9 xenopus lae
42	107	7.3	525	11	Q80VR9	Q80vr9 mus musculus
43	107	7.3	543	4	Q8IU72	Q8iu72 homo sapien
44	107	7.3	627	4	Q7Z328	Q7z328 homo sapien
45	106	7.3	567	3	Q9UUS7	Q9uus7 collettotric

#### ALIGNMENTS

#### RESULT 1

Q9UL42 PRELIMINARY; PRT; 283 AA.  
 AC Q9UL42;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Paraneoplastic cancer-testis-brain antigen (Fragment).  
 GN MA4  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J.,  
 RA Posner J.B., Dalmay J.;  
 RT "Identification of a novel cancer testis brain antigen using serum  
 RT antibodies from patients with testicular tumors and paraneoplastic  
 RT limbic encephalomyelitis."  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF083115; AA05626.1; -;  
 DR InterPro; IPR005162; Retrotrans\_gag.  
 DR Pfam; PF03732; Retrotrans\_gag; 1.  
 FT NON TER 1  
 SQ SEQUENCE 283 AA; 32333 MW; E27BA6BCDDCD240A4 CRC64;

Query Match 100.0%; Score 1462; DB 4; Length 283;  
 Best Local Similarity 100.0%; Pred. No. 3,8e-107;  
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VQGGGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALQEGVSPATVPCISPEL 60  
 Db 1 VQGGGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALQEGVSPATVPCISPEL 60  
 Oy 61 LAHLGQMAHAPOQLLPMPYRKLRFVSGSAVPAPEEESFEVWLEQATEIVKEWPVTEAE 120  
 Db 61 LAHLGQMAHAPOQLLPMPYRKLRFVSGSAVPAPEEESFEVWLEQATEIVKEWPVTEAE 120

QY 121 KRWLAESLRGPDALDMLHIVQADNPSISVEECLEAFKQVFGSLGSRRTAQRVYLKTYQEE 180  
 DB |||||||  
 QY 121 KRWLAESLRGPDALDMLHIVQADNPSISVEECLEAFKQVFGSLGSRRTAQRVYLKTYQEE 180  
 DB |||||||  
 QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCRRLRELKDGQPP 240  
 DB |||||||  
 QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCRRLRELKDGQPP 240  
 DB |||||||

QY 241 PSFLELMKVIREEEESAFENESIEEPEERDGYGRWNHEGDD 283  
 DB |||||||  
 QY 241 PSFLELMKVIREEEESAFENESIEEPEERDGYGRWNHEGDD 283  
 DB |||||||

RESULT 2  
 O94959 ID O94959 PRELIMINARY; PRT; 364 AA.  
 AC O94959;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein KIAA0883 (Paraneoplastic associated brain-testis-  
 cancer antigen).  
 DE KIAA0883.  
 GN KIAA0883.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99156230; PubMed=10048485;  
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 5:355-364(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Eichen J.G., Dalmay J., Wade D., Rosenfeld M.R.;  
 RT "Characterization of a Brain-Testis-Cancer Antigen.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF206590; BAA74906.1; -  
 DR EMBL; AF286487; AAC28165.1; -  
 DR Genew; HGNC:9159; PNM2.  
 DR InterPro; IPR005162; Retrotrans\_gag.  
 DR Pfam; PF03732; Retrotrans\_gag; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0E93 CRC64;

Query Match 99.8%; Score 1459; DB 4; Length 364;  
 Best Local Similarity 99.8%; Pred. No. 9.1e-107; Indels 0; Gaps 0;  
 Matches 282; Conservative 1; Mismatches 0;  
 QY 1 VQKGKGVKVIKFTPNQDTFLERLNLFLEKEGQTVSGMFRALGQEGVSPATVPCISPEL 60  
 DB |||||||  
 QY 82 VQKGKGVKVIKFTPNQDTFLERLNLFLEKEGQTVSGMFRALGQEGVSPATVPCISPEL 141  
 DB |||||||  
 QY 61 LAHLGQAMAHAPQPLLPYRKLRVFGSAVPAPPEESFEVWLEQATEIVKWPVTEAE 120  
 DB |||||||  
 QY 142 LAHLGQAMAHAPQPLLPYRKLRVFGSAVPAPPEESFEVWLEQATEIVKWPVTEAE 201  
 DB |||||||  
 QY 121 KRWLAESLRGPDALDMLHIVQADNPSISVEECLEAFKQVFGSLGSRRTAQRVYLKTYQEE 180  
 DB |||||||  
 QY 202 KRWLAESLRGPDALDMLHIVQADNPSISVEECLEAFKQVFGSLGSRRTAQRVYLKTYQEE 261  
 DB |||||||  
 QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCRRLRELKDGQPP 240  
 DB |||||||  
 QY 262 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCRRLRELKDGQPP 321  
 DB |||||||

Query Match 99.8%; Score 1459; DB 4; Length 364;  
 Best Local Similarity 99.8%; Pred. No. 9.1e-107; Indels 0; Gaps 0;  
 Matches 282; Conservative 1; Mismatches 0;  
 QY 1 VQKGKGVKVIKFTPNQDTFLERLNLFLEKEGQTVSGMFRALGQEGVSPATVPCISPEL 60  
 DB |||||||  
 QY 82 VQKGKGVKVIKFTPNQDTFLERLNLFLEKEGQTVSGMFRALGQEGVSPATVPCISPEL 141  
 DB |||||||  
 QY 61 LAHLGQAMAHAPQPLLPYRKLRVFGSAVPAPPEESFEVWLEQATEIVKWPVTEAE 120  
 DB |||||||  
 QY 142 LAHLGQAMAHAPQPLLPYRKLRVFGSAVPAPPEESFEVWLEQATEIVKWPVTEAE 201  
 DB |||||||  
 QY 121 KRWLAESLRGPDALDMLHIVQADNPSISVEECLEAFKQVFGSLGSRRTAQRVYLKTYQEE 180  
 DB |||||||  
 QY 202 KRWLAESLRGPDALDMLHIVQADNPSISVEECLEAFKQVFGSLGSRRTAQRVYLKTYQEE 261  
 DB |||||||  
 QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCRRLRELKDGQPP 240  
 DB |||||||  
 QY 262 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCRRLRELKDGQPP 321  
 DB |||||||  
 QY 241 PSFLELMKVIREEEESAFENESIEEPEERDGYGRWNHEGDD 283  
 DB |||||||  
 QY 322 PSFLELMKVIREEEESAFENESIEEPEERDGYGRWNHEGDD 364  
 DB |||||||

## RESULT 3

O9GMU3 ID O9GMU3 PRELIMINARY; PRT; 364 AA.  
 AC O9GMU3;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 DE Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OC NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AB047632; BAB12156.1; -  
 DR InterPro; IPR005162; Retrotrans\_gag.  
 DR Pfam; PF03732; Retrotrans\_gag; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;

Query Match 97.7%; Score 1428; DB 6; Length 364;  
 Best Local Similarity 97.2%; Pred. No. 2.5e-104; Indels 0; Gaps 0;  
 Matches 275; Conservative 4; Mismatches 4;  
 QY 1 VQKGKGVKVIKFTPNQDTFLERLNLFLEKEGQTVSGMFRALGQEGVSPATVPCISPEL 60  
 DB |||||||  
 QY 82 VQKGKGVKVIKFTPNQDTFLERLNLFLEKEGQTVSGMFRALGQEGVSPATVPCISPEL 141  
 DB |||||||  
 QY 61 LAHLGQAMAHAPQPLLPYRKLRVFGSAVPAPPEESFEVWLEQATEIVKWPVTEAE 120  
 DB |||||||  
 QY 142 LAHLGQAMAHAPQPLLPYRKLRVFGSAVPAPPEESFEVWLEQATEIVKWPVTEAE 201  
 DB |||||||  
 QY 121 KRWLAESLRGPDALDMLHIVQADNPSISVEECLEAFKQVFGSLGSRRTAQRVYLKTYQEE 180  
 DB |||||||  
 QY 202 KRWLAESLRGPDALDMLHIVQADNPSISVEECLEAFKQVFGSLGSRRTAQRVYLKTYQEE 261  
 DB |||||||  
 QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCRRLRELKDGQPP 240  
 DB |||||||  
 QY 262 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCRRLRELKDGQPP 321  
 DB |||||||  
 QY 241 PSFLELMKVIREEEESAFENESIEEPEERDGYGRWNHEGDD 283  
 DB |||||||  
 QY 322 PSFLELMKVIREEEESAFENESIEEPEERDGYGRWNHEGDD 364  
 DB |||||||

## RESULT 4

O8BHKO ID O8BHKO PRELIMINARY; PRT; 365 AA.  
 AC O8BHKO;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to KIAA0883 protein.  
 GN PNM2 OR A830049P17RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cortex;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;

"Analysis of the mouse transcriptome based on functional annotation of

QY	195	LRKAVEKRAIPRRITADQVRLEQVWAGATLNQMLWCRLRELKDQGPSPFLELMKVIREEE	254
Db	61	LRRAVEXRAIPRRITADQVRLEQVWAGATLNQMLWCRLRELKDQGPSPFLELMKVIREEE	120
QY	255	EEEAASFNESTIEEPFERDGYGRWNHEGDD	283
Db	121	EEEAASFNESTIEEPFERDGYGRWNHEGDD	149
RESULT 6			
Q8VHZ4			
ID	Q8VHZ4	PRELIMINARY;	PRT; 353 AA.
AC	Q8VHZ4;		
DT	01-MAR-2002 (TrEMBLrel. 20, Created)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Paraneoplastic onconeural protein MAL.		
GN	MAL.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI TaxID=10116		

RT	"Analysis of the mouse transcriptome based on functional annotation of			
ET	60,770 full-length cDNAs."			
EL	Nature 420:565-573(2002).			
DR	EMSL; AK043718; BAC31626.1; -			
DR	EMSL; AK043910; BAC31700.1; -			
DR	MGD; MGI:2444129; Pma2.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0005215; F:transporter activity; IEA.			
DR	GO; GO:0006810; P:transport; IEA.			
DR	InterPro; IPR000425; MIP.			
DR	InterPro; IPR005162; Retrotrans gag.			
DR	Pfam; PF03732; Retrotrans_gag_1.			
DR	PROSITE; PS00221; MIP; 1			
SQ	SEQUENCE 365 AA; 41201 MW; 0F90C940B9D843D9 CRC64;			
Query Match 81.6%; Score 1193.5; DB 11; Length 365;				
Best Local Similarity 82.1%; Pred.No. 7.3e-86;				
Matches 230; Conservative 21; Mismatches 28; Indels 1; Gaps 1;				

RA Poaner J.B., Rosenfeld M.R.;  
 RT "Ma1, a novel neuron- and testis-specific protein, is recognized by  
 RT the serum of patients with paraneoplastic neurological disorders.";  
 RL Brain 122:27-39(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum;  
 RA Dalmau J., Rosenfeld M.R., Voltz R., Hoard R.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF037364; AAD13810.3; -;  
 DR Genew; HGNC:9158; PNMAL.  
 DR GO; GO:0005737; C:cyttoplasm; TAS.  
 DR GO; GO:0005730; C:nucleolus; TAS.  
 DR GO; GO:0007417; P:central nervous system development; TAS.  
 DR GO; GO:0007283; P:spermatogenesis; TAS.  
 DR GO; GO:0007283; P:spermatogenesis; TAS.  
 SQ SEQUENCE 353 AA; 39800 MW; 3B841691AB89AD3D CRC64;  
 Query Match 43.0%; Score 628; DB 4; Length 353;  
 Best Local Similarity 47.9%; Pred. No. 2.4e-41;  
 Matches 128; Conservative 52; Mismatches 75; Indels 12; Gaps 4;  
 QY 3 GKGWVKVIFKTPNQDTEFLERLNLFLKKGQTVSGMFRALQGVSPATVPCISPELLA 62  
 Db 83 GKGWVKVLFKPTSDAEFLERLHLFLAREGWTVDVARVLGFQNPPTTP-----GPEMPA 138  
 QY 63 HLLQAMAHAPQLL-PMRYKLRVFGSAPVAPPEESFEVWLEQATEIVKEMPVTEAK 121  
 Db 139 EMLNYLDNVIOPLVESIWKRLTLFSGRDIPOGGEETFDPMLEHTNEVLEEWQVSDVEK 198  
 QY 122 KRWLAESLRGPAALDLMHIVQADNPISVBECELEAFKQVFGSLESRRTAQVRYLKYQEG 181  
 Db 199 RRLMESLRGPAADVIRILKSNPAITTAECLEKALEQVFGSVESRDQIKELNTYQNP 258  
 QY 182 EKVSAYVLRLETLRKAVKRAIPRIADQVRLEQVMAGA-----TLNQMLWCLRELK 237  
 Db 259 EKLSAYVLRLEPLQKVKVKGAKDKNVQARLEQVIAGANHSGAIRQLWL---TGAGE 315  
 QY 238 GPPSFLELMKVIREEEERASFENES 264  
 Db 316 GPAPNLQOLLVQIREEAKKEEAEAA 342  
 RESULT 8  
 Q8NG07 PRELIMINARY; PRT; 353 AA.  
 ID Q8NG07  
 AC Q8NG07; 22, Created)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DE Paraneoplastic antigen.  
 GN PNMAL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Schueller M.M., Jenne D.E., Schutze dit Belkner N., Hohlfield R.,  
 RA Voltz R.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Prostate;  
 RA Strausberg R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF320308; AAN05100.1; -;  
 DR EMBL; BC039577; AAK39577.1; -;  
 DR EMBL; BC039577; AAK39577.1; -;  
 SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;  
 Query Match 43.0%; Score 628; DB 4; Length 353;  
 Best Local Similarity 47.9%; Pred. No. 2.4e-41;  
 Matches 128; Conservative 52; Mismatches 75; Indels 12; Gaps 4;

QY 3 GKGWVKVIFKTPNQDTEFLERLNLFLKKGQTVSGMFRALQGVSPATVPCISPELLA 62  
 Db 83 GKGWVKVLFKPTSDAEFLERLHLFLAREGWTVDVARVLGFQNPPTTP-----GPEMPA 138  
 QY 63 HLLQAMAHAPQLL-PMRYKLRVFGSAPVAPPEESFEVWLEQATEIVKEMPVTEAK 121  
 Db 139 EMLNYLDNVIOPLVESIWKRLTLFSGRDIPOGGEETFDPMLEHTNEVLEEWQVSDVEK 198  
 QY 122 KRWLAESLRGPAALDLMHIVQADNPISVBECELEAFKQVFGSLESRRTAQVRYLKYQEG 181  
 Db 199 RRLMESLRGPAADVIRILKSNPAITTAECLEKALEQVFGSVESRDQIKELNTYQNP 258  
 QY 182 EKVSAYVLRLETLRKAVKRAIPRIADQVRLEQVMAGA-----TLNQMLWCLRELK 237  
 Db 259 EKLSAYVLRLEPLQKVKVKGAKDKNVQARLEQVIAGANHSGAIRQLWL---TGAGE 315  
 QY 238 GPPSFLELMKVIREEEERASFENES 264  
 Db 316 GPAPNLQOLLVQIREEAKKEEAEAA 342  
 RESULT 9  
 Q8UZW8 PRELIMINARY; PRT; 466 AA.  
 ID Q8UZW8  
 AC Q8UZW8; 22, Created)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Similar to paraneoplastic antigen MA3.  
 GN PNMAL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RA Strausberg R.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC036726; AAK36726.1; -;  
 DR MGD; MGI:2180565; Pma3.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001544; Aminotrans IV.  
 DR InterPro; IPR001878; Znf CCHC.  
 DR Pfam; PF00098; zf-CCHC; 1  
 DR PRINTS; PR00939; C2HCZNFINGER.  
 DR PROSITE; P550158; ZF\_CCHC; 1.  
 DR PROSITE; P550158; ZF\_CCHC; 1.  
 SQ SEQUENCE 466 AA; 54040 MW; 113787E37B0B3AAD CRC64;  
 Query Match 42.5%; Score 622; DB 11; Length 466;  
 Best Local Similarity 49.4%; Pred. No. 1e-40;  
 Matches 132; Conservative 55; Mismatches 72; Indels 8; Gaps 4;  
 QY 1 VQGGGKGVKVIKTPNQDTEFLERLNLFLKKGQTVSGMFRALQGVSPATVPCISPE- 59  
 Db 82 IEKGKGPWEVWVKKPHSDDEFLNHLNHLFLEERRTVSDMNRVLGTHSNHPTKTTISADF 141  
 QY 60 -LLAHLIGQMAHAPQLL-PMRYKLRVFGSAPVAPPEESFEVWLEQATEIVKWPVT 117  
 Db 142 WYWAQTLGAVM----QPLLEQMLYRELVRVFGNTISIFGLLAFDSWLEHTTEMLQWQVP 197  
 QY 118 EAKKEWLAEALRGPAALDLMHIVQADNPISVBECELEAFKQVFGSLESRRTAQVRYLKY 177  
 Db 198 EVEKRRRLMECLRGPAALQVNVVLRANNAITVKECLEALQVFGSVDRNRKIAQLKPKAY 257  
 QY 178 QEEGKVSAYVLRLETLRKAVKRAIPRIADQVRLEQVMAGATLNQMLWCLRELK 237  
 Db 258 QEPGEKVSFVRLLETLQKALEKNAISRKNVQNLRLKRLIGGAILSAKLREKLKMLKOR 317  
 QY 238 GPPSFLELMKVIREEEERASFENE 263







Blank

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:35:22 ; Search time 52.4349 Seconds  
(without alignments)  
1194.261 Million cell updates/sec

Title: US-10-037-860-7  
Perfect score: 996  
Sequence: 1 PLALLEDWCRMSVDEQKSL.....EESFEVWLEQATEIVKWP 195

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	996	100.0	120	13	US-10-037-860-7
2	604	60.6	125	12	US-09-804-014A-42
3	593	59.5	283	13	US-10-037-860-11
4	466	46.8	353	9	US-09-965-529-7
5	466	46.8	353	10	US-09-969-680A-7
6	462.5	46.4	463	13	US-10-037-860-13
7	441	44.3	329	13	US-10-037-860-4
8	440	44.2	318	12	US-09-804-014A-40
9	436	43.8	312	12	US-09-804-014A-73
10	436	43.8	312	12	US-09-804-014A-74
11	436	43.8	321	12	US-09-804-014A-39
12	436	43.8	351	9	US-09-965-529-1
13	436	43.8	351	10	US-09-969-680A-1
14	436	43.8	351	12	US-09-804-014A-16
15	436	43.8	351	15	US-10-341-434-10

16	434	43.6	452	16	US-10-408-765A-2385
17	345	34.6	399	15	US-10-094-749-1978
18	327.5	32.9	204	14	US-10-029-386-33747
19	306	30.7	120	12	US-09-804-014A-41
20	270	27.1	116	9	US-09-864-761-34645
21	146	14.7	538	16	US-10-408-765A-2392
22	137.5	13.8	584	12	US-10-221-278-355
23	137.5	13.8	584	15	US-10-291-172-355
24	99.5	10.0	403	15	US-10-094-466-38
25	92.5	9.3	337	12	US-10-296-115-1208
26	89	8.9	342	13	US-10-001-857-201
27	87.5	8.8	255	12	US-10-087-192-213
28	87.5	8.8	311	12	US-09-727-100-1
29	87.5	8.8	1357	15	US-10-295-027-1199
30	87.5	8.8	3859	16	US-10-408-765A-354
31	86	8.6	1083	15	US-10-369-493-4443
32	86	8.6	1083	15	US-10-369-493-7202
33	86	8.6	1084	12	US-10-282-122A-49912
34	85.5	8.6	407	15	US-10-369-493-17903
35	84.5	8.5	520	9	US-09-213-678-2
36	84.5	8.5	520	14	US-10-032-585-7035
37	84	8.4	336	9	US-09-745-763-17
38	84	8.4	1638	13	US-10-090-458-2
39	84	8.4	1642	13	US-10-090-458-5
40	84	8.4	1642	14	US-10-005-338B-5
41	84	8.4	3613	14	US-10-156-761-10432
42	83.5	8.4	492	9	US-09-764-898-206
43	82	8.2	408	12	US-10-425-114-55513
44	82	8.2	547	16	US-10-437-963-199553
45	82	8.2	607	12	US-10-424-599-218036

## ALIGNMENTS

RESULT 1  
US-10-037-860-7  
; Sequence 7, Application US/10037860  
; Publication No. US20020123114A1  
; GENERAL INFORMATION:  
; APPLICANT: Jerome B. Posner  
; APPLICANT: Joseph O. Dalmay  
; APPLICANT: Myrina R. Rosenfeld  
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma  
; TITLE OF INVENTION: ANTIBODIES  
; FILE REFERENCE: 2581.1004-004  
; CURRENT APPLICATION NUMBER: US/10/037,860  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: 09/189,527  
; PRIOR FILING DATE: 1998-11-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-037-860-7

Query Match 100.0%; Score 996; DB 13; Length 195;  
Best Local Similarity 100.0%; Pred. No. 2.8e-95;  
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	PLALLEDWCRMSVDEQKSLMTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQEN	60
Db	1	PLALLEDWCRMSVDEQKSLMTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQEN	60
Qy	61	ANAVLLELLETDVSAIPSEVQGGVKKVIFKTPNODTEFLERINLFLKEGQTVSGMF	120
Db	61	ANAVLLELLETDVSAIPSEVQGGVKKVIFKTPNODTEFLERINLFLKEGQTVSGMF	120
Qy	121	RALGQEAISPAIVPCISPELLAHLGQMAHAPOLPLPMRYRKLRFVSGSAVPAPEESF	180
Db	121	RALGQEAISPAIVPCISPELLAHLGQMAHAPOLPLPMRYRKLRFVSGSAVPAPEESF	180

Wed Sep 22 12:06:48 2004

us-10-037-860-7.rapb

```

; LENGTH: 283
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-11

Query Match      59.5%; Score 593; DB 13; Length 283;
Best Local Similarity 98.3%; Pred. No. 4.6e-53;
Matches 113; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 81 VQKGGVWVKVFKTPNQDTFELRLNLFLEKQGVSGMFRALQGEALSPATVPCISP 140
DB 1 VQKGGVWVKVFKTPNQDTFELRLNLFLEKQGVSGMFRALQGEALSPATVPCISP 60
;
QY 141 LAHLQGAHAHAPQPLPMRYKRLRVFSGSAVPAPPEESFEVWLEQATEIVKEWP 195
DB 61 LAHLQGAHAHAPQPLPMRYKRLRVFSGSAVPAPPEESFEVWLEQATEIVKEWP 115
;

RESULT 4
US-09-965-529-7
; Sequence 7, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyoung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 2483172CD1
US-09-965-529-7

Query Match      46.8%; Score 466; DB 9; Length 353;
Best Local Similarity 47.9%; Pred. No. 1.1e-39;
Matches 93; Conservative 38; Mismatches 57; Indels 6; Gaps 4;

QY 2 LALLEDCRIMSVDEQKSLMTGIPADFEAEIQEVLTSLGRYLLGKIFRKQENA 61
DB 3 MTLLEDCRGMVDVNSQRALLVWGIPVNCDEAEIETLQAAMPQVS-YRWLGRMFWRENA 61
;
QY 62 NAVLELLEDDTDSAIPSEVQKGGVWVKVFKTPNQDTFELRLNLFLEKQGVSGMFR 121
DB 62 KAALLELTCAVDYAAIPREMPCKGKGVWVKVLPKPTSDAEFLERLHLFLAREGTVQDVAR 121
;
QY 122 ALGQEAALSPATVPCISPPELLAHLQGAHAHAPQPL-PMRYKRLRVFSGSAVPAPPEESF 180
DB 122 VLGFQ--NPTPTP--GPEMPAEMLYILDNVIQPLVESIWKYKRLTFLSGRDIPGGEFTF 177
;
QY 181 EVWLEQATEIVKEW 194
DB 178 DPWLEHTNEVLEW 191
;

RESULT 5
US-09-969-680A-7
; Sequence 7, Application US/09969680A

; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-42

Query Match      60.6%; Score 604; DB 12; Length 120;
Best Local Similarity 100.0%; Pred. No. 9.5e-55;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LALLEDCRIMSVDEQKSLMTGIPADFEAEIQEVLTSLGRYLLGKIFRKQENA 61
DB 1 LALLEDCRIMSVDEQKSLMTGIPADFEAEIQEVLTSLGRYLLGKIFRKQENA 60
;
QY 62 NAVLELLEDDTDSAIPSEVQKGGVWVKVFKTPNQDTFELRLNLFLEKQGVSGMFR 121
DB 61 NAVLELLEDDTDSAIPSEVQKGGVWVKVFKTPNQDTFELRLNLFLEKQGVSGMFR 120
;

RESULT 3
US-10-037-860-11
; Sequence 11, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerone B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
```

[illegible]

```
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 40
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (20)
; OTHER INFORMATION: Wherin Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification
; US-09-804-014A-40

Query Match 44.2%; Score 440; DB 12; Length 318;
Best Local Similarity 46.4%; Pred. No. 4.7e-37;
Matches 90; Conservative 36; Mismatches 62; Indels 6; Gaps 4;

Qy 2 LALLEDCRIMSVDEQKSLMTGTPADFEAEIOEVLQETLKSIGRYRLLGKIFRKQENA 61
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 MTLLEDWCRGMDVNSQKLLVWGLPVNCDIAEIEETLQAAFPQVS-YMLGRMFWRENA 61
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 62 NAVLELLEDTDVSAIPSEVQGGVWKVIFKTPNQDTEFLERLNLFLEKEGQTVSGMFR 121
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 KAALLELTGAVYAAIPREMPGKGGVWKVFKPPTSDAEFLERLHLFLAREGWTQDVAR 121
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 122 ALGQEALSPATVPCISPELLAHLLGOAHAPOLLP-PMRYRKLRFVSGSAVPAPEESF 180
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 VLGFQ--NPTTP--GPMPAEMLYILDNVLIQVLESIIWYKRLTLFLSGKGHPRAWRGNF 177
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 181 EVMLEQATEIVKKEW 194
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 DPWLEHTNEVLEEW 191
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-09-804-014A-73
; Sequence 73, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 74
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-804-014A-74
```

```
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 73
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-804-014A-73

Query Match 43.8%; Score 436; DB 12; Length 312;
Best Local Similarity 46.9%; Pred. No. 1.2e-36;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

Qy 2 LALLEDCRIMSVDEQKSLMTGTPADFEAEIOEVLQETLKSIGRYRLLGKIFRKQENA 61
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LRLLDWCRCMDMNPRAKLLIAGISQSCSVAEIEEALQAGLADPLGEYRLLRMPFRDENR 62
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 62 NAVLELLEDTDVSAIPSEVQGGVWKVIFKTPNQDTEFLERLNLFLEKEGQTVSGMFR 121
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 KVALVGLTAEETSHALVPEIKPGKGIWRVIFKPPDPDNTFLSRNEFLAGEGWTVGELSR 122
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 122 ALGQEALSPATVPCISPELLAHLLGOAHAPOLLP-MRYRKLRFVSGSAVPAPEESF 180
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 ALGHENGSLDPQGMPEMMAPLAQAAL-EALQPALOCLKYLKLVFSGRESPPGSEEF 181
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 181 EVMLEQATEIVKKEW 194
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GRWMEHTTQMKAW 195
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-09-804-014A-74
; Sequence 74, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 74
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-804-014A-74
```

```

Query Match      43.8%; Score 436; DB 9; Length 351;
Best Local Similarity 46.9%; Pred. No. 1.4e-36;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2

QY 2 LALLEDCWRIMSVDEQKSLMVTGIPADFEAEATQEVLOETUKSLGRYRLLLGKIFRKQENA 61
Db 1 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
3 LRULEDCWRGDMWNPRAKLIAGISQSCSVAETEEALQAGLAPLGEYRLLLGRMPFRDENR 62
QY 62 NAVILLELTDVSATPSEVQGGVGMVKFVKTPNODTEFLERLNLFLKEGQTVSGMFR 121
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPDNTFLSLRNEFLAGEGNTVGEUSR 122
QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPOLLP-MRYRKLVFSGSAVPAPEBSF 180
Db 1 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
123 ALCHENGSLDPQCGMTPENWAPMLAQL-EALQPALQCLYKXKLVFSGRESPGGEHF 181
QY 181 EYWLEQATEIVKEW 194
Db 1 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
182 GRWMFHTTQMIKAW 195

RESULT 13
US-09-969-680A-1
; Sequence 1, Application US/09969680A
; Publication NO. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14

```

Prior Application Number: 60/149,641  
Prior Filing Date: 1999-08-17  
Prior Application Number: 60/164,203  
Prior Filing Date: 1999-11-09  
Number of Seq ID NOS: 74  
Software: PERL Program  
Seq ID No 1  
Length: 351  
Type: PRT  
Organism: Homo sapiens  
Feature:  
Name/Key: misc feature  
Other Information: Incyte ID No. US20030124649A1 112301CD1  
US-09-969-680A-1

Query Match 43.8%; Score 436; DB 10; Length 351;  
Best Local Similarity 46.9%; Pred. No. 1.4e-36;  
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

Qy 2 LALLEDCRIMSVDEQKSLMVTGIPADFEAEIQTETLSLGRYRLGKIPRKQENA 61  
Db 3 LRLLEDWCRGMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGEYRLGMRFRDENR 62

Qy 62 NAVILLELDDTDSVAIPSEVQGGVWVKVIFKTPNQDTFELERLNLFLKEGQTVSGMFR 121  
Db 63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTFLSRLNEFLAGEGTVGELSR 122

Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPPEESF 180  
Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQALQCLKYKKLRVFSGRESPEPGESEF 181

Qy 181 EVWLEQATEIVKEW 194  
Db 182 GRWMEHTTQMIKAW 195

RESULT 15  
US-10-341-434-10  
Sequence 10, Application US/10341434  
Publication No. US20030215835A1  
GENERAL INFORMATION:  
APPLICANT: Origene Technologies  
TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes  
FILE REFERENCE: 9U 204 205 R1  
CURRENT APPLICATION NUMBER: US/10/341,434  
CURRENT FILING DATE: 2003-07-18  
PRIOR APPLICATION NUMBER: US 60/348,164  
PRIOR FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: US 60/348,119  
PRIOR FILING DATE: 2002-01-15  
NUMBER OF SEQ ID NOS: 238  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 10  
LENGTH: 351  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-341-434-10

Query Match 43.8%; Score 436; DB 15; Length 351;  
Best Local Similarity 46.9%; Pred. No. 1.4e-36;  
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

Qy 2 LALLEDCRIMSVDEQKSLMVTGIPADFEAEIQTETLSLGRYRLGKIPRKQENA 61  
Db 3 LRLLEDWCRGMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGEYRLGMRFRDENR 62

Qy 62 NAVILLELDDTDSVAIPSEVQGGVWVKVIFKTPNQDTFELERLNLFLKEGQTVSGMFR 121  
Db 63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTFLSRLNEFLAGEGTVGELSR 122

Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPPEESF 180  
Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQALQCLKYKKLRVFSGRESPEPGESEF 181

Qy 181 EVWLEQATEIVKEW 194  
Db 182 GRWMEHTTQMIKAW 195

Search completed: September 21, 2004, 14:04:07  
Job time : 53.4349 secs

Prior Application Number: 60/149,641  
Prior Filing Date: 1999-08-17  
Prior Application Number: 60/164,203  
Prior Filing Date: 1999-11-09  
Number of Seq ID NOS: 74  
Software: PERL Program  
Seq ID No 1  
Length: 351  
Type: PRT  
Organism: Homo sapiens  
Feature:  
Name/Key: misc feature  
Other Information: Incyte ID No. US20030124649A1 112301CD1  
US-09-969-680A-1

Query Match 43.8%; Score 436; DB 10; Length 351;  
Best Local Similarity 46.9%; Pred. No. 1.4e-36;  
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

Qy 2 LALLEDCRIMSVDEQKSLMVTGIPADFEAEIQTETLSLGRYRLGKIPRKQENA 61  
Db 3 LRLLEDWCRGMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGEYRLGMRFRDENR 62

Qy 62 NAVILLELDDTDSVAIPSEVQGGVWVKVIFKTPNQDTFELERLNLFLKEGQTVSGMFR 121  
Db 63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTFLSRLNEFLAGEGTVGELSR 122

Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPPEESF 180  
Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQALQCLKYKKLRVFSGRESPEPGESEF 181

Qy 181 EVWLEQATEIVKEW 194  
Db 182 GRWMEHTTQMIKAW 195

RESULT 14  
US-09-804-014A-16  
Sequence 16, Application US/09804014A  
Publication No. US20030064489A1  
GENERAL INFORMATION:  
APPLICANT: Li, Li  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Vernet, Corine  
APPLICANT: Fernandes, Elma  
APPLICANT: Shinkets, Richard  
APPLICANT: Spaderna, Steven  
APPLICANT: Majumder, Kumud  
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same  
FILE REFERENCE: 15966-721 US  
CURRENT APPLICATION NUMBER: US/09/804,014A  
CURRENT FILING DATE: 2002-04-24  
PRIOR APPLICATION NUMBER: 60/188,316  
PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: 60/188,277  
PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: 60/189,139  
PRIOR FILING DATE: 2000-03-14  
PRIOR APPLICATION NUMBER: 60/189,140  
PRIOR FILING DATE: 2000-03-14  
PRIOR APPLICATION NUMBER: 60/190,401  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/190,231  
PRIOR FILING DATE: 2000-03-17  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 16  
LENGTH: 351  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-804-014A-16

Query Match 43.8%; Score 436; DB 12; Length 351;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 ; Search time 36.0037 Seconds  
(without alignments)  
1530.308 Million cell updates/sec

Title: US-10-037-860-7

Perfect score: 996

Sequence: 1 PLALLEDWCRIMSVDEQKSL.....EESEFVWLEQATEIVKEWP 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	996	100.0	195	3	AAB12526	Aab12526 Human Ma2
2	982	98.6	364	7	ADC08977	Adc08977 Onconeuro
3	590	59.2	283	3	AAB12528	Aab12528 Human Ma4
4	466	46.8	353	4	AAB74701	Aab74701 Human mem
5	462.5	46.4	455	5	ABB05727	Abb05727 Human s19
6	462.5	46.4	463	3	AAB12529	Aab12529 Human Ma5
7	462.5	46.4	463	3	AAB42315	Aab42315 Human ORF
8	436	43.8	312	3	AAB43023	Aab43023 Human ORF
9	436	43.8	351	4	AAE01340	Aae01340 Human gen
10	436	43.8	351	4	AAB74695	Aab74695 Human mem
11	436	43.8	351	4	AAU08664	Aau08664 Human NOV
12	435	43.7	329	3	AAAB12525	Aab12525 Human Ma1
13	434	43.6	452	6	AAO16179	Aao16179 Human pro
14	345	34.6	399	6	ADA54410	Ada54410 Human pro
15	345	34.6	399	6	ABG99947	Abg99947 Human nov
16	342	34.3	439	6	ABP75736	Abp75736 Human sec
17	270	27.1	116	4	AAAB14937	Aam14937 Peptide #
18	270	27.1	116	4	ABB33906	Abb33906 Peptide #
19	270	27.1	116	4	AAAM27367	Aam27367 Peptide #
20	270	27.1	116	4	ABB28727	Abb28727 Peptide #
21	270	27.1	116	4	ABB19347	Abb19347 Protein #
22	270	27.1	116	4	AAAM67075	Aam67075 Human bon
23	270	27.1	116	4	AAAM54673	Aam54673 Human bra
24	270	27.1	116	4	ABG48741	Abg48741 Human liv
25	270	27.1	116	4	AAAM02665	Aam02665 Peptide #

#### ALIGNMENTS

##### RESULT 1

AAB12526  
ID AAB12526 standard; protein; 195 AA.

XX AAB12526;

XX AC AAB12526;

XX DT 02-NOV-2000 (first entry)

XX DE Human Ma2 protein SEQ ID NO:7.

XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;  
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;  
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;  
KW germ-cell tumour.

XX OS Homo sapiens.

XX PN JP2000146982-A.

XX PD 26-MAY-2000.

XX PF 10-NOV-1999; 99JP-00320171.

XX PR 10-NOV-1998; 98US-00189527.

XX PA (SLOK ) SLOAN KETTERING INST CANCER RES.

XX DR WPI; 2000-468119/41.

XX N-PSDB; AAA60834.

XX PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic  
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test  
PT sample for the presence or absence of antibodies to a Ma family  
PT polypeptide.

XX Claim 48; Fig 2; 27pp; Japanese.

XX The present invention describes a method for diagnosing a paraneoplastic  
CC syndrome or neoplasm. The method comprises assessing a test sample for  
CC the presence or absence of antibodies to a Ma family polypeptide (I). The  
CC method is used to diagnose a paraneoplastic syndrome especially  
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or  
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing  
CC antibodies to (i) preferably Ma1, which is indicative presence of breast  
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular  
CC cancer, germ-cell tumours or Ma2, which is indicative of testicular  
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the  
CC Ma2 protein as given in the present invention

Abg36734 Human pep  
Aae01336 Human gen  
Abu52641 Human bra  
Adc31124 Human nov  
Abg19651 Novel hum  
Abg14259 Novel hum  
Aab94854 Human pro  
Aau28186 Novel hum  
Abg97495 Human NOV  
Aam51624 KIAA0883 -  
Abo14772 Novel hum  
Abo14773 Novel hum  
Aab60478 Human cel  
Aam5693 Human pro  
Abg13816 Novel hum  
Ade08994 Novel pro  
Abg12037 Novel hum  
Abg05411 Novel hum  
Abp52958 Human lun  
Aar43684 Lipopolys

26 270 27.1 116 5 ABG36734  
27 160 16.1 280 4 AAE01336  
28 146 14.7 538 4 ABU52641  
29 146 14.7 538 7 ADC31124  
30 144 14.5 615 4 ABG19651  
31 144 14.5 615 4 ABG14259  
32 142 14.3 237 4 AAB94854  
33 137.5 13.8 584 4 AAU28186  
34 99.5 10.0 403 5 ABG97495  
35 99.5 10.0 403 5 AAM51624  
36 99.5 10.0 403 6 ABO14772  
37 99.5 10.0 403 6 ABO14773  
38 97.5 9.8 402 4 AAB60478  
39 92.5 9.3 337 4 AAM25693  
40 92 9.2 872 4 ABG13816  
41 92 9.2 872 7 ADE08994  
42 90 9.0 521 4 ABG12037  
43 90 9.0 521 4 ABG05411  
44 89 8.9 342 5 ABP52958  
45 87.5 8.8 311 2 AAR43684

XX SQ Sequence 195 AA;  
 Query Match 100.0%; Score 996; DB 3; Length 195;  
 Best Local Similarity 100.0%; Pred. NO. 2.1e-96;  
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLALLEDDWCRIMSVDEQKSLMTGTIPADFEAEIQVLOETLKSGLRYRLGKIFRKQEN 60  
 DB 1 PLALLEDDWCRIMSVDEQKSLMTGTIPADFEAEIQVLOETLKSGLRYRLGKIFRKQEN 60

QY 61 ANAVLLELLEDDTVSAIPSEVQKGGVWVKVIFKTPNQDTFELRLNLFLEKEGQTVSGMF 120  
 DB 61 ANAVLLELLEDDTVSAIPSEVQKGGVWVKVIFKTPNQDTFELRLNLFLEKEGQTVSGMF 120

QY 121 RALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYRKLRFVSGSAVPAPEESF 180  
 DB 121 RALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYRKLRFVSGSAVPAPEESF 180

QY 181 EVMLEQATEIVKEMP 195  
 DB 181 EVMLEQATEIVKEMP 195

RESULT 2  
 ADC08977  
 ID ADC08977 standard; protein; 364 AA.  
 AC ADC08977;  
 DT 18-DEC-2003 (first entry)  
 XX Onconeural antigen Ma2 protein.  
 DE Human; Ma2; onconeural; antigen; Alzheimer's disease;  
 KW neurodegenerative disease; diagnosis; neuroprotective; gene therapy.  
 XX Homo sapiens.  
 OS  
 XX W02003073104-A2.  
 XX  
 PD 04-SEP-2003.  
 XX  
 PF 26-FEB-2003; 2003WO-EP001946.  
 XX  
 PR 26-FEB-2002; 2002EP-00004177.  
 PR 26-FEB-2002; 2002US-0359307P.  
 XX  
 PA (EVOT-) EVOTEC NEUROSCIENCES GMBH.  
 XX  
 PI Hipfel R, Von Der Kammer H, Pohlner J;  
 XX  
 DR WPI; 2003-721818/68.  
 DR GENBANK; O94959, KIAA0883.  
 XX  
 PT Diagnosing or prognosticating, or determining increased risk of  
 PT developing a neurodegenerative disease by determining level or activity  
 PT of a transcription or translation product of a gene coding for Ma  
 PT onconeural antigen.  
 XX  
 PS Disclosure; Fig 9; 51pp; English.  
 XX  
 CC The present sequence is the protein sequence of human onconeural  
 CC antigen Ma2. The invention discloses the detection and differential  
 CC expression and regulation of the Ma2 gene in specific brain regions of AD  
 CC patients. The Ma2 gene and its transcription and/or translation products  
 CC may have a causative role in the regional selective neuronal degeneration  
 CC typically observed in AD, or may confer a neuroprotective function to the  
 CC remaining nerve cells. Methods are claimed for diagnosing or  
 CC prognosticating a neurodegenerative disease, for monitoring the  
 CC progression of a neurodegenerative disease, and for evaluating treatment  
 CC of a neurodegenerative disease, especially AD, in a subject by  
 CC determining the level and/or activity of a transcription or translation

CC product of an Ma onconeural antigen gene, especially Ma2. Also claimed  
 CC are: a method for treating or preventing AD and related neurodegenerative  
 CC disorders using the Ma2 gene or its transcription or translation product;  
 CC a method of screening for modulating agents of neurodegenerative diseases  
 CC ; and a recombinant non-human animal comprising an Ma2 gene sequence,  
 CC which is useful for screening, testing and validating candidate  
 CC diagnostic and therapeutic agents.  
 XX Sequence 364 AA;  
 SQ

Query Match 98.6%; Score 982; DB 7; Length 364;  
 Best Local Similarity 99.0%; Pred. NO. 1.6e-94;  
 Matches 192; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LALLEDDWCRIMSVDEQKSLMTGTIPADFEAEIQVLOETLKSGLRYRLGKIFRKQENA 61  
 DB 3 LALLEDDWCRIMSVDEQKSLMTGTIPADFEAEIQVLOETLKSGLRYRLGKIFRKQENA 62

QY 62 NAVLLELLEDDTVSAIPSEVQKGGVWVKVIFKTPNQDTFELRLNLFLEKEGQTVSGMF 121  
 DB 63 NAVLLELLEDDTVSAIPSEVQKGGVWVKVIFKTPNQDTFELRLNLFLEKEGQTVSGMF 122

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYRKLRFVSGSAVPAPEESF 181  
 DB 123 ALGQEGVSPATVPCISPELLAHLGQAMAHAPQPLLPMRYRKLRFVSGSAVPAPEESF 182

QY 182 VMLEQATEIVKEMP 195  
 DB 183 VMLEQATEIVKEMP 196

RESULT 3  
 AAB12528  
 ID AAB12528 standard; protein; 283 AA.  
 XX  
 AC AAB12528;  
 XX  
 DT 02-NOV-2000 (first entry)  
 XX  
 DE Human Ma4 protein SEQ ID NO:11.  
 XX  
 KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;  
 KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;  
 KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;  
 KW germ-cell tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP2000146982-A.  
 XX  
 PD 26-MAY-2000.  
 XX  
 PF 10-NOV-1999; 99JP-00320171.  
 XX  
 PR 10-NOV-1998; 98US-00189527.  
 XX  
 PA (SLOK ) SLOAN KETTERING INST CANCER RES.  
 XX  
 XX WPI; 2000-468119/41.  
 XX N-PSDB; AAA60836.  
 XX  
 PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic  
 PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test  
 PT sample for the presence or absence of antibodies to a Ma family  
 PT polypeptide.  
 XX  
 CC Claim 48; Fig 7-8; 27pp; Japanese.  
 CC  
 CC The present invention describes a method for diagnosing a paraneoplastic  
 CC syndrome or neoplasm. The method comprises assessing a test sample for  
 CC the presence or absence of antibodies to a Ma family polypeptide (1). The  
 CC method is used to diagnose a paraneoplastic syndrome especially  
 CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or

CC neoplasm in an individual. The method diagnoses the neoplasm by assessing  
 CC antibodies to (I) preferably Ma1, which is indicative presence of breast  
 CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular  
 CC cancer and germ-cell tumours or Ma2, which is indicative of testicular  
 CC cancer, germ-cell tumour, and lung cancer. The present sequence is the  
 CC Ma4 protein as given in the present invention  
 XX  
 SQ

Query Match 59.2%; Score 590; DB 3; Length 283;  
 Best Local Similarity 97.4%; Pred. No. 2.3e-53;  
 Matches 112; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 81 VGGKGVKVIPTNQDTFELRLNLFLEKGGTVSGMFRALGOEALSPATVPCISPEL 140  
 DB 1 VGGKGVKVIPTNQDTFELRLNLFLEKGGTVSGMFRALGOEALSPATVPCISPEL 60  
 QY 141 LAHLGQAMAHAPQLLPMRYKRLRVFSGSAVPAPEESFEVWLEQATEIVKEMP 195  
 DB 61 LAHLGQAMAHAPQLLPMRYKRLRVFSGSAVPAPEEDSFEVWLEQATEIVKEMP 115

RESULT 4  
 AAB74701  
 ID AAB74701 standard; protein; 353 AA.  
 AC AAB74701;  
 XX

DT 12-JUN-2001 (first entry)

DE Human membrane associated protein MEMAP-7.

KW Human; membrane associated protein; MEMAP; diagnosis; cytostatic;  
 KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;  
 KW antiarteriosclerotic; gene therapy; cell proliferative disorder;  
 KW autoimmune disorder; inflammatory disorder; neurological disorder;  
 KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;  
 KW epilepsy; diarrhoea.  
 XX

OS Homo sapiens.

FN WO200112662-A2.

PD 22-FEB-2001.

PF 14-AUG-2000; 2000WO-US022315.

PR 17-AUG-1999; 99US-0149641P.

PR 09-NOV-1999; 99US-0164203P.

PA (INCY-) INCYTE GENOMICS INC.

PI Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;

PI Baughn MR, Lu DAM, Patterson C;

DR N-PSDB; AAF81747.

DR WPI; 2001-168860/17.

XX Isolated polypeptide with a human membrane associated protein sequence is  
 PT useful for the diagnosis, prevention and treatment of cell proliferative,  
 PT autoimmune/inflammatory, neurological and gastrointestinal disorders.  
 XX

PS Claim 1; Page 119-120; 173pp; English.

XX AAF81741 to AAF81777 encode the human membrane associated proteins  
 CC (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,  
 CC antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and  
 CC antiarteriosclerotic activities, which can be used in gene therapy.  
 CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition  
 CC associated with decreased expression of functional MEMAP and antagonists  
 CC of MEMAP are used to treat a disease or condition associated with  
 CC overexpression of functional MEMAP. These disorders include cell  
 CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal

CC disorders. The MEMAP polynucleotides and proteins are also used for the  
 CC diagnosis of these disorders. Specific examples of these disorders  
 CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.  
 CC MEMAP proteins can be used to screen for compounds which specifically  
 CC bind MEMAP including antibodies, oligonucleotides, proteins and small  
 CC molecules. MEMAP polynucleotides can be used to prepare transgenic  
 CC animals which can be studied to provide information concerning human  
 CC disease. Anti-MEMAP antibodies are useful in immunoassays for the  
 CC detection of MEMAP protein and can be used as antagonists to treat or  
 CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP  
 CC can be delivered to target cells with genetic abnormalities with respect  
 CC to the expression of MEMAP to treat or prevent a disorder associated with  
 CC MEMAP

XX Sequence 353 AA;

Query Match 46.8%; Score 466; DB 4; Length 353;

Best Local Similarity 47.9%; Pred. No. 3.7e-40;

Matches 93; Conservative 38; Mismatches 57; Indels 6; Gaps 4;

QY 2 LALLDWCGRMSVDQKSLMYTGIPADFEAEIQEVLTSLGRLYLLGKIFRKOENA 61

DB 3 MTLLEDWCGRMDVNSQRALLVWGIPVNCDEABIEBTLQAMPQVS-YRMLGRMFREENA 61

QY 62 NAVLLELLEDDTVSAIPSEVOGKGVKVIPTNQDTFELRLNLFLEKGGTVSGMFR 121

DB 62 KAALLELTCAVDYAAIPREMPGKGVKVIPTNQDTFELRLNLFLEKGGTVSGMFR 121

QY 122 ALGOEALSPATVPCISPELLAHLGQAMAHAPQLL-PMRYKRLRVFSGSAVPAPEESF 180

DB 122 VLGFQ--NPTPTP--GPMPAEMNLVILDNVIQPLVESIWKYKRLTFLFSGRDIPGGETTF 177

QY 181 EVWLEQATEIVKEW 194

DB 178 DPWLEHTNEVLEEW 191

RESULT 5

ABB05727

ID ABB05727 standard; protein; 455 AA.

AC ABB05727;

DT 30-APR-2002 (first entry)

DE Human signal transduction protein clone tes3\_5k22.

XX Human; foetal brain; foetal kidney; melanoma; testis; amygdala;

KW gene therapy.

OS Homo sapiens.

XX WO2001198454-A2.

XX 27-DEC-2001.

PF 25-APR-2001; 2001WO-IB002050.

PR 25-APR-2000; 2000US-0199380P.

XX (GEHU-) GERMAN HUMAN GENOME PROJECT.

PI Wiemann S;

XX WPI; 2002-055860/07.

DR N-PSDB; ABA93764.

XX Human cDNA sequences and clones derived from human fetal brain, fetal  
 PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic  
 PT screening and therapy.

PS Claim 1; Page 382; 611pp; English.

CC The present invention describes assemblages and computer readable media  
 CC comprising novel human cDNA sequences and clones derived from human  
 CC foetal brain, foetal kidney, melanoma, testis and anygdala cDNA  
 CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the  
 CC present invention which encode the proteins given in ABB05662 to  
 CC ABB05729. The human cDNA sequences and clones can be used in gene  
 CC therapy. The clones may be used in a variety of applications, for example  
 CC they may be used in profiling assays, for providing large arrays of human  
 CC genetic material for implementing large-scale screening strategies and  
 CC for treating diseases via gene therapy procedures

XX  
 SQ Sequence 455 AA;  
 Query Match 46.4%; Score 462.5; DB 5; Length 455;  
 Best Local Similarity 47.7%; Pred. No. 1.3e-39;  
 Matches 93; Conservative 34; Mismatches 65; Indels 3; Gaps 2;  
 QY 1 PLALLEDWCRIMSVDRQKSLMTVGIPADFEABIQEVLQETLKSGLRYRLGKIFRQEN 60  
 DB 2 PLTLQDWCRCGEHLNTRCMLILGIPEDCGEDEFEEFTLQACRHLGRYRVIGMFRREN 61  
 QY 61 ANAVILLELLEDDTVSAIPSEVQKGGVWKFVFTPNQDTFELRLNLFLEKEGQTVSGMF 120  
 DB 62 AQAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERETVSDMN 121  
 QY 121 RALGQEAALSPATVPCISPELLAHLGQAMAHAPQPLL-PMRYKRLRVFSGSAVPAPEES 179  
 DB 122 RVLGSDTNCSPARTVTSPEFWT--WAQTGLGAQVQPLLEQMLYRELRFVSGNTISIPGALA 179  
 QY 180 FEVWLEQATEIVKWEK 194  
 DB 180 FDAWLEHTTEMLQMW 194

RESULT 6  
 AAB12529  
 ID AAB12529 standard; protein; 463 AA.

XX AAB12529;  
 XX  
 DT 02-NOV-2000 (first entry)  
 XX  
 DE Human Ma5 protein SEQ ID NO:13.  
 XX  
 KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;  
 KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;  
 KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;  
 KW germ-cell tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP2000146982-A.  
 XX  
 PD 26-MAY-2000.  
 XX  
 PF 10-NOV-1999; 99JP-00320171.  
 XX  
 PR 10-NOV-1998; 98US-00189527.  
 XX  
 PA (SLOK ) SLOAN KETTERING INST CANCER RES.  
 XX  
 DR WPI; 2000-468119/41.  
 DR N-PSDB; AAA60837.  
 XX  
 PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic  
 PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test  
 PT sample for the presence or absence of antibodies to a Ma family  
 PT polypeptide.  
 XX  
 PS Claim 48; Fig 9-10; 27pp; Japanese.  
 XX  
 CC The present invention describes a method for diagnosing a paraneoplastic  
 CC syndrome or neoplasm. The method comprises assessing a test sample for

CC the presence or absence of antibodies to a Ma family polypeptide (I). The  
 CC method is used to diagnose a paraneoplastic syndrome especially  
 CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or  
 CC neoplasm in an individual. The method diagnoses the neoplasm by assessing  
 CC antibodies to (i) preferably Ma1, which is indicative presence of breast  
 CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular  
 CC cancer, germ-cell tumours or Ma2, which is indicative of testicular  
 CC cancer, germ-cell tumour, and lung cancer. The present sequence is the  
 CC Ma5 protein as given in the present invention

XX  
 SQ Sequence 463 AA;  
 Query Match 46.4%; Score 462.5; DB 3; Length 463;  
 Best Local Similarity 47.7%; Pred. No. 1.3e-39;  
 Matches 93; Conservative 34; Mismatches 65; Indels 3; Gaps 2;  
 QY 1 PLALLEDWCRIMSVDRQKSLMTVGIPADFEABIQEVLQETLKSGLRYRLGKIFRQEN 60  
 DB 2 PLTLQDWCRCGEHLNTRCMLILGIPEDCGEDEFEEFTLQACRHLGRYRVIGMFRREN 61  
 QY 61 ANAVILLELLEDDTVSAIPSEVQKGGVWKFVFTPNQDTFELRLNLFLEKEGQTVSGMF 120  
 DB 62 AQAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERETVSDMN 121  
 QY 121 RALGQEAALSPATVPCISPELLAHLGQAMAHAPQPLL-PMRYKRLRVFSGSAVPAPEES 179  
 DB 122 RVLGSDTNCSPARTVTSPEFWT--WAQTGLGAQVQPLLEQMLYRELRFVSGNTISIPGALA 179  
 QY 180 FEVWLEQATEIVKWEK 194  
 DB 180 FDAWLEHTTEMLQMW 194

RESULT 7  
 AAB42315  
 ID AAB42315 standard; protein; 463 AA.

XX AAB42315;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF2079 polypeptide sequence SEQ ID NO:4158.  
 XX  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW immunosuppressant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antineoplastic;  
 KW antiviral; antibacterial; antifungal; antineoplastic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antineoplastic disease; coagulation;  
 KW thrombosis; contraceptive.

OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US008621.  
 XX  
 PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX  
 CC (CURA-) CURAGEN CORP.

PI Shimkets RA, Leach M;  
 XX WPI; 2000-602362/57.  
 DR N-PSDB; AAC76524.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 11; Page 3345-3347; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 SQ Sequence 463 AA;  
 Query Match 46.4%; Score 462.5; DB 3; Length 463;  
 Best Local Similarity 47.7%; Pred. No. 1.3e-39;  
 Matches 93; Conservative 34; Mismatches 65; Indels 3; Gaps 2;  
 QY 1 PLALLEDCRIMSVDEQKSLMVTGTPADFEAEIOEVLTQKSLGRYLLGKIFRKQEN 60  
 Db 2 PLTLQDWCRCHEHLNTRCMLGILPEDCGEDEFETLQACRHLYRVRVTRGRMFRREN 61  
 QY 61 ANAVILLELLEDTVSAIPSEVQGGKGVKWKVFKTPNODTEFLERLNLFLEKEGQTVSGMF 120  
 Db 62 AQAILLELAQDIDVALLPREIPGKGPWEVIVKPNSDGFEFLNRLNLEERRVTSVMN 121  
 QY 121 RALQGEALSPATVPCISPELLAHLGQAWAHAPQLP-PMRYKLRVFGSVAVPAPEES 179  
 Db 122 RVLGSDTNCSPRVVTISPEFWT--WAQTLGAAVQPLLEQMLYRELVRVFGNTISIPGALA 179  
 QY 180 FEVWLEQATEIVKEW 194  
 Db 180 FDWLEHTTEMLQKW 194

cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 cholesterol ester storage; systemic lupus erythematosus; infection;  
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 thrombosis; contraceptive.  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US008621.  
 XX  
 PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX  
 PA (CUPA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Leach M;  
 XX  
 WPI; 2000-602362/57.  
 N-PSDB; AAC77232.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 11; Page 4759; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 SQ Sequence 312 AA;  
 Query Match 43.8%; Score 436; DB 3; Length 312;  
 Best Local Similarity 46.9%; Pred. No. 4.5e-37;  
 Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;  
 QY 2 LALLEDCRIMSVDEQKSLMVTGTPADFEAEIOEVLTQKSLGRYLLGKIFRKQENA 61  
 Db 3 LRLLEDWCRCMDMPKALLTAGISQSCVAEIEEALQAGLAPLGEYLLGRMFRREN 62  
 QY 62 NAVILLELLEDTVSAIPSEVQGGKGVKWKVFKTPNODTEFLERLNLFLEKEGQTVSGMF 121  
 Db 63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTFLSRNLNEFLAGEMTVGLSR 122  
 QY 122 ALQGEALSPATVPCISPELLAHLGQAWAHAPQLP-PMRYKLRVFGSVAVPAPEES 180  
 Db 123 ALGHENGSLDPEQGMIPEMWAPMLAQL-EALQALQCLKYKLRVFGSPSPGEEF 181  
 QY 181 EVWLEQATEIVKEW 194

CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The  
CC present sequence represents a human secreted protein fragment referred to  
CC in the disclosure of the invention

XX SQ Sequence 351 AA;

Query Match 43.8%; Score 436; DB 4; Length 351;  
Best Local Similarity 46.9%; Pred. No. 5.4e-37;  
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2

Qy 2 LALLEDCRIMSVDEQKSLMVTGIPADFEAEIQEVLQETILKSGLRYLLGIKIFRKQENA 61  
Db | :  
3 LRLEDDCRGMNPNRAALLIAGISQSQAIEEALQAGLAPLGEYRLLGEMFRDNR 62

Qy 62 NAVLLELEDTSVAIPSEVQGCGVKVIPTNQDTFELRLNLFLKEGGQTIVSGMFR 121  
Db | :  
63 KVALVGLTATSTSHALVPKEIPKGKIWRVFKPPDPDNFTLSRLNEFLAGGMTVGELSR 122

Qy 122 ALGOEALS PATVPCISPELLAHLIGQAMAHAPQLLP-MRYRKLRFSGSAVPAPEEESF 180  
Db | :  
123 ALHGENSELDPEQMIPEMWAPMLAQAL-EALQPALQCLKYKKLRVTFSGRESPEGESEEF 181

Qy 181 EVMLEQATEIVKEW 194  
Db | :  
182 GRWMFHTTOMIKAW 195

RESULT 10  
AAB74695  
ID AAB74695 standard; protein; 351 AA.  
XX AAB74695;  
XX AC  
XX DT  
XX DE 12-JUN-2001 (first entry)  
XX Human membrane associated protein MEMAP-1.  
XX Human; membrane associated protein; MEMAP; diagnosis; cytostatic;  
KW anti-inflammatory; anticonvulsant; immunosuppressive; anti-diarrheic;  
KW anti-arteriosclerotic; gene therapy; cell proliferative disorder;  
KW autoimmune disorder; inflammatory disorder; neurological disorder;  
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;  
KW epilepsy; diarrhoea.  
XX OS Homo sapiens.  
XX WO200112662-A2.  
XX PN  
XX PD 22-FEB-2001.  
XX PF 14-AUG-2000; 2000WO-US022315.  
XX PR 17-AUG-1999; 99US-0149641P.  
XX PR 09-NOV-1999; 99US-0164203P.  
XX FA (INCY-) INCYTE GENOMICS INC.  
XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;  
PI Baughn MR, Lu DAM, Patterson C;  
XX WPI; 2001-168860/17.  
DR N-PADB; AAF81741.  
XX Isolated polypeptide with a human membrane associated protein sequence is  
PT useful for the diagnosis, prevention and treatment of cell proliferative,  
PT autoimmune/inflammatory, neurological and gastrointestinal disorders.  
XX Claim 1; Page 114-115; 173pp; English.

XX AAF81741 to AAF81777 encode the human membrane associated proteins  
CC (MEMAP) given in AAB74695 to AAB74731. MEMAPS have cytostatic,  
CC antiinflammatory, anticonvulsant, immunosuppressive, anti-diarrheic and  
CC antiarteriosclerotic activities, which can be used in gene therapy.



```

DT 02-NOV-2000 (first entry)
XX
DE Human Mal protein SEQ ID NO:4.
XX
KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
KW germ-cell tumour.
XX
OS Homo sapiens.
XX
PN JP2000146982-A.
XX
PD 26-MAY-2000.
XX
PF 10-NOV-1999; 99JP-00320171.
XX
PR 10-NOV-1998; 98US-00189527.
XX
PA (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
XX WPI; 2000-468119/41.
XX N-PSDB; AAA60833.
XX
PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
PT sample for the presence or absence of antibodies to a Ma family
PT polypeptide.
XX
XX Claim 48; Fig 1; 27pp; Japanese.
XX
CC The present invention describes a method for diagnosing a paraneoplastic
CC syndrome or neoplasm. The method comprises assessing a test sample for
CC the presence or absence of antibodies to a Ma family polypeptide (I). The
CC method is used to diagnose a paraneoplastic syndrome especially
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
CC antibodies to (I) preferably Ma1, which is indicative presence of breast
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
CC cancer and germ-cell tumours or Ma2, which is indicative of testicular
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
CC Ma1 protein as given in the present invention
XX
SQ Sequence 329 AA;
Query Match 43.7%; Score 435; DB 3; Length 329;
Best Local Similarity 45.9%; Pred. No. 6.2e-37; Indels 6; Gaps 4;
Matches 89; Conservative 37; Mismatches 62;
QY 2 LALLEDCWCRIMSDVEQKSLMVTGIPADFEAEIQLQETLKSIGRYRLGKIFRKQENA 61
DQ 3 MTLLEDWCRGMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWEENA 61
QY 62 NAVLELLEDVTSAPSEVQGGVWVKVIFKTPNQDTEFLERINLFLKEGQTVSGMFR 121
DQ 62 KAALLELTGAVDYAAIPREMPGKGWVKVLFKPTSDAEFLERLHLLEAREGWTVDVAR 121
QY 122 ALGQEALSPATVPCI-SPELLAHLIGQAMAHAPQPLL-PMRYKLRVFGSSAVPAPERESF 180
DQ 122 VLGFQ--NPTPTP--GPEMPAEMLNYLDNIQPLVESIWKYKTLFLSGKHPRWNGNF 177
QY 181 EVMLEQATEIVKWEW 194
DQ 178 DPWLEHTNEVLEEW 191
RESULT 13
AAO16179
ID AAO16179 standard; protein; 452 AA.
XX
AC AAO16179;
XX
XX 28-MAR-2003 (first entry)
DE

```

---

```

XX
DE Human protein #5.
XX
KW Human; vaccine; adult whole brain; foetal whole brain; tonsil;
KW adult hippocampus; disease-associated SNP analysis; knockout mouse;
KW disease model mouse; cancer; neurological disorder.
XX
OS Homo sapiens.
XX
PN WO200299103-A1.
XX
PD 12-DEC-2002.
XX
PF 27-MAY-2002; 2002WO-JP005134.
XX
PR 04-JUN-2001; 2001JP-00168370.
XX
PR 16-AUG-2001; 2001JP-00246915.
XX
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX N-PSDB; AAL51207.
XX
XX Ohara O, Nagase T, Nakajima D;
XX WPI; 2003-140622/13.
XX N-PSDB; AAL51207.
XX
PT DNA preferentially expressed in human adult and fetal brain tissue useful
PT for diagnosis, treatment and analysis of cancer and mental disorders.
XX
XX Claim 1; Page 56-60; 73pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of seven
CC human proteins that are preferentially expressed in adult whole brain,
CC foetal whole brain, tonsil and adult hippocampus tissue. The DNA
CC sequences are useful for the analysis of disease-associated single
CC nucleotide polymorphisms and the production of knockout and human disease
CC model mice. The DNA and protein sequences of the invention are useful for
CC the prevention (vaccine) and treatment of cancer and neurological
CC disorders. The present amino acid sequence represents a human protein of
CC the invention
XX
SQ Sequence 452 AA;
Query Match 43.6%; Score 434; DB 6; Length 452;
Best Local Similarity 44.2%; Pred. No. 1.3e-36;
Matches 88; Conservative 39; Mismatches 54; Indels 18; Gaps 3;
QY 2 LALLEDCWCRIMSDVEQKSLMVTGIPADFEAEIQLQETLKSIGRYRLGKIFRKQENA 61
DQ 7 LTLLEDWCRGMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWEENA 66
QY 62 NAVLELLEDVTSAPSEVQGGVWVKVIFKTPNQDTEFLERINLFLKEGQTVSGMFR 121
DQ 67 KAVFIELADVTNTTLEPSHIPGKGSEWVVKPRNPDEFLSRNLNLYFLKDEGRSMTDVAR 126
QY 122 ALG-----QBALSPATVPCI-SPELLAHLIGQAMAHAPQPLL-PMRYKLRVFGSSAVPAP 175
DQ 127 ALGCCSLPAESLDAEVMPPQVRSPL-----EPPKESMWYKLRVFGSGTASPSF 174
QY 176 EERSFEVLEQATEIVKWEW 194
DQ 175 GEFTFEDWLEQVTEIMPIW 193
RESULT 14
ADA54410
ID ADA54410 standard; protein; 399 AA.
XX
AC ADA54410;
XX
XX 20-NOV-2003 (first entry)
XX
XX Human protein, SEQ ID 1978.
DE

```

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
 KW inflammatory disease; osteoporosis; neurological disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1293569-A2.  
 XX  
 PD 19-MAR-2003.  
 XX  
 PF 21-MAR-2002; 2002EP-00006586.  
 XX  
 PR 14-SEP-2001; 2001JP-00328381.  
 PR 24-JAN-2002; 2002US-0350435P.  
 XX  
 PA (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX  
 DR WPI; 2003-395539/38.  
 DR N-PSDB; ADA52771.  
 XX  
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.  
 XX  
 PS Claim 14; SEQ ID NO 1978; 205pp; English.  
 XX  
 CC The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 XX  
 SQ Sequence 399 AA;  
 Query Match 34.6%; Score 345; DB 6; Length 399;  
 Best Local Similarity 38.9%; Pred. No. 2.6e-27;  
 Matches 77; Conservative 38; Mismatches 63; Indels 20; Gaps 5;  
 QY 2 LALLDWCRCIMSDVQKSLMTGTPADFEAEIQVLQETLSGLRYLLGKIPKQENA 61  
 DB 3 VTMLQDWCRCWGNVARRGLLIGIPEDCDDEAFQSLAALRPMGHFTVLGKAFREEDNA 62  
 QY 62 NAVLELLEDTDVSAIPSEVQKGGVWVIFKTPNQDTEF--LERINLFLEKEGQTVSGM 119  
 DB 63 TAALVELDREVNIALVPREIPGTGPGWNVFVPRCSGEFLGLGRVTFHPEEGQMVESV 122  
 QY 120 FRALGOEALSPATVPCISPPELLAHL--LQGMAMHAPQLL-PMRYRKURVFGSGAVPAPE 176  
 DB 123 AGALG-----VGLRRVCWLRISGQAV----QPWVEAVRCQSLGLVFGSGRDPAPG 167  
 QY 177 EESFEVWLEQATEIVKEW 194  
 DB 168 EESFEVWLDHTTEMLHVV 185  
 RESULT 15  
 ABG99947  
 ID ABG99947 standard; protein; 399 AA.  
 AC ABG99947;  
 XX  
 DT 17-JAN-2003 (first entry)  
 XX  
 DE Human novel polypeptide #60.  
 XX  
 KW Human; genetic disorder; gene mapping; medical imaging; cancer;  
 KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;

KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;  
 KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;  
 KW fungal infection; bacterial infection; autoimmune disease; diabetes;  
 KW atopic dermatitis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200274961-A1.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PF 14-MAR-2002; 2002WO-US005109.  
 XX  
 PR 15-MAR-2001; 2001US-00810173.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;  
 XX  
 DR WPI; 2003-040556/03.  
 DR N-PSDB; ABX05045.  
 XX  
 PT New isolated polypeptides and polynucleotides, useful for preventing,  
 PT treating or ameliorating medical conditions, such as cancer,  
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative  
 PT disorders, and infections.  
 XX  
 PS Claim 9; SEQ ID NO 586; 235pp; English.  
 XX  
 CC The invention relates to human polynucleotides and the polypeptides they  
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,  
 CC forensics, gene mapping, medical imaging, identification of mutations,  
 CC responsible for genetic disorders or other traits, assessing biodiversity  
 CC and producing many other types of data and products dependent on DNA and  
 CC amino acid sequences. They are also useful for preventing, treating or  
 CC ameliorating medical conditions, such as cancer, neurodegenerative  
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell  
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,  
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or  
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).  
 CC Sequences ABG99888-ABG99889 and ABU00010-ABU00433 represent human  
 CC polypeptides of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied by the European Patent Office  
 XX  
 SQ Sequence 399 AA;  
 Query Match 34.6%; Score 345; DB 6; Length 399;  
 Best Local Similarity 38.9%; Pred. No. 2.6e-27;  
 Matches 77; Conservative 38; Mismatches 63; Indels 20; Gaps 5;  
 QY 2 LALLDWCRCIMSDVQKSLMTGTPADFEAEIQVLQETLSGLRYLLGKIPKQENA 61  
 DB 3 VTMLQDWCRCWGNVARRGLLIGIPEDCDDEAFQSLAALRPMGHFTVLGKAFREEDNA 62  
 QY 62 NAVLELLEDTDVSAIPSEVQKGGVWVIFKTPNQDTEF--LERINLFLEKEGQTVSGM 119  
 DB 63 TAALVELDREVNIALVPREIPGTGPGWNVFVPRCSGEFLGLGRVTFHPEEGQMVESV 122  
 QY 120 FRALGOEALSPATVPCISPPELLAHL--LQGMAMHAPQLL-PMRYRKURVFGSGAVPAPE 176  
 DB 123 AGALG-----VGLRRVCWLRISGQAV----QPWVEAVRCQSLGLVFGSGRDPAPG 167  
 QY 177 EESFEVWLEQATEIVKEW 194  
 DB 168 EESFEVWLDHTTEMLHVV 185  
 Search completed: September 21, 2004, 13:35:13  
 Job time : 37.0037 secs

B/ANK

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 45.1859 Seconds  
(without alignments)  
1361.621 Million cell updates/sec

Title: US-10-037-860-7

Perfect score: 996

Sequence: 1 PLALLEDWCIRMSVDQKSL.....EEESFEVWLEQATEIVKWP 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	982	98.6	364	4	O94959
2	972	97.6	194	4	O95145
3	959	96.3	364	6	Q9GMU3
4	766	76.9	365	11	Q8BHK0
5	593	59.5	283	4	Q9UL42
6	468	47.0	353	11	Q8VHZ4
7	466	46.8	353	4	Q8NG07
8	465	46.7	353	4	Q95144
9	462.5	46.4	455	4	Q9H0A4
10	462.5	46.4	463	4	Q9UL41
11	462	46.4	353	11	Q9CYP2
12	459.5	46.1	466	11	Q8JZW8
13	459	46.1	353	11	Q8CIC8
14	436	43.8	351	4	Q96BY2
15	436	43.8	351	4	Q9HAS1
16	434	43.6	448	4	Q8NET3

17	434	43.6	452	4	Q96PV4
18	431	43.3	351	6	Q95K14
19	431	43.3	352	11	Q9ERH6
20	345	34.6	399	4	Q96A40
21	338.5	34.0	246	11	Q8C533
22	338.5	34.0	430	11	Q80VM8
23	337	33.8	378	4	Q8N3H4
24	337	33.8	435	4	Q86V59
25	211.5	21.2	327	11	Q8VC32
26	146	14.7	538	4	Q9H0W5
27	146	14.7	538	4	Q8TB26
28	142	14.3	237	4	Q9H833
29	107.5	10.8	402	11	Q9CZA5
30	102	10.2	608	2	Q68043
31	100.5	10.1	393	11	Q9DB17
32	100.5	10.1	393	11	Q8VD24
33	99.5	10.0	403	4	Q8TE36
34	97.5	9.8	402	4	Q8N1C1
35	95	9.5	341	16	Q9WV83
36	92	9.2	192	4	Q8ND90
37	91	9.1	455	16	Q7V8D5
38	87.5	8.8	229	11	Q8R1J2
39	87.5	8.8	269	16	Q88LH9
40	87.5	8.8	311	11	Q8VE43
41	87.5	8.8	897	10	Q7X6G9
42	87.5	8.8	1684	4	Q75218
43	87.5	8.8	3830	4	Q9Y6H4
44	87.5	8.8	3859	4	Q9Y631
45	86.5	8.7	959	16	Q7USU6

#### ALIGNMENTS

RESULT 1

O94959 PRELIMINARY; PRT; 364 AA.  
AC O94959;  
DT 01-MAY-1999 (TREMblrel. 10, Created)  
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
DE Hypothetical protein KIAA0883 (Paraneoplastic associated brain-testis-cancer antigen).  
DN KIAA0883.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99156230; PubMed=10048485;  
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirotsawa M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";  
RL DNA Res. 5:355-364 (1998).  
RN [2]  
SEQUENCE FROM N.A.  
RA Eichen J.G., Dalmay J., Wade D., Rosenfeld M.R.;  
RT "Characterization of a Brain-Testis-Cancer Antigen.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF286487; BAA74906.1; -.  
DR EMBL; AF286487; AAG28165.1; -.  
DR Genew; HGNC:9159; PNMA2.  
DR InterPro: IPR005162; Retrotrans gag.  
DR Pfam; PF03732; Retrotrans gag; I.  
KW Hypothetical protein.  
SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96B3F0E93 CRC64;

Query Match 98.6%; Score 982; DB 4; Length 364;  
Best Local Similarity 99.0%; Pred. No. 7.1e-81;

Matches 192; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 61  
 Db 3 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 62  
 Qy 62 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121  
 Db 63 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 122  
 Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESFE 181  
 Db 123 ALGQEGVSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESFE 182  
 Qy 182 VWLEQATEIVKEWP 195  
 Db 183 VWLEQATEIVKEWP 196

RESULT 2  
 Q95145  
 ID O95145 PRELIMINARY; PRT; 194 AA.  
 AC O95145;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE Paraneoplastic neuronal antigen NM2 (Fragment).  
 GN NM2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum;  
 RA Dalmat J., Rosenfeld M.R., Voltz R., Hoard R.;  
 RT "antigen recognized by serum of patients with paraneoplastic  
 RT neurologic disorders."  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF037365; AAD02098.1; --  
 FT NON\_TER  
 SQ SEQUENCE 194 AA; 21808 MW; 2854E4D5BED3F28C CRC64;

Query Match 97.6%; Score 972; DB 4; Length 194;  
 Best Local Similarity 97.9%; Pred. No. 2.6e-80;  
 Matches 190; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 61  
 Db 1 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 60  
 Qy 62 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121  
 Db 61 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 120  
 Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESFE 181  
 Db 121 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESFE 180  
 Qy 182 VWLEQATEIVKEWP 195  
 Db 181 VWLEQATEIVKEWP 194

RESULT 3  
 Q9GMU3  
 ID Q9GMU3 PRELIMINARY; PRT; 364 AA.  
 AC Q9GMU3;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB047632; BAB12156.1; --  
 DR InterPro; IPR005162; Retrotrans\_gag.  
 DR Pfam; PF03732; Retrotrans\_gag; 1.  
 DR Hypothetical protein.  
 SQ SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;

Query Match 96.3%; Score 959; DB 6; Length 364;  
 Best Local Similarity 95.9%; Pred. No. 8.8e-79;  
 Matches 186; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 61  
 Db 3 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 62  
 Qy 62 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121  
 Db 63 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 122  
 Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESFE 181  
 Db 123 ALGHEGMSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESFE 182  
 Qy 182 VWLEQATEIVKEWP 195  
 Db 183 VWLEQATEIVKEWP 196

RESULT 4  
 Q8BHK0  
 ID Q8BHK0 PRELIMINARY; PRT; 365 AA.  
 AC Q8BHK0;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to KIAA0883 protein.  
 GN FNMA2 OR A830049P17RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cortex;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573 (2002).  
 DR EMBL; AK043718; BAC31626.1; --  
 DR EMBL; AK043910; BAC31700.1; --  
 DR MGD; MGI:2444129; Pnma2.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000425; MIP.  
 DR InterPro; IPR005162; Retrotrans\_gag.  
 DR Pfam; PF03732; Retrotrans\_gag; 1.  
 DR PROSITE; PS00221; MIP; 1.  
 SQ SEQUENCE 365 AA; 41201 MW; 0F90C940B9D843D9 CRC64;



US Homo sapiens (human) .  
OC Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK028331; BAC25885.1; -.
DR MGD; MGI:2180564; Pnuma.
SQ SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;

Query Match 46.1%; Score 459; DB 11; Length 353;
Best Local Similarity 47.9%; Pred. No. 2.3e-33;
Matches 93; Conservative 36; Mismatches 59; Indels 6; Gaps 4;

QY 2 LALLEDCRIMSVDEQKSLMTGIPADFEAEIQTQVLQETLKSIGRYLLGKIFRKOENA 61
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 MTLLEDWCRCMDVNSORALLVWGPVNCDETEIETLQAAMPQVS-YRVILGRMFWRENA 61
||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 NAVLELLEDTDVSAIPSEVQGGVWKVFKTPNQDTFLERLNLFLKEGQTVSGMFR 121
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 KAALLELTGAVDSLIPREMPGKGLWKVFKPTSDAVFLERHLFLAREGTVQDVAR 121
||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLL-PMRYRKLRFVSGSAVPAPPEESF 180
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 VLGQ--NPAP--GPETPAEMLNYLDNVQPLVESIWYKXLTLPFGKDIPGGBETF 177
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 EVMLEQATEIVKWEW 194
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 DSMLEHSENIIEEW 191
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q96BY2 PRELIMINARY; PRT; 351 AA.
AC Q96BY2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Modulator of apoptosis 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC015044; AAH15044.1; -.
DR Genew; HGNC:16658; MOAP1.
SQ SEQUENCE 351 AA; 39512 MW; 5310142AC02B563C CRC64;

Query Match 43.8%; Score 436; DB 4; Length 351;
Best Local Similarity 46.9%; Pred. No. 2.9e-31;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

QY 2 LALLEDCRIMSVDEQKSLMTGIPADFEAEIQTQVLQETLKSIGRYLLGKIFRKOENA 61
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LRLLEDWCRCMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGETYLLGRMFRDNR 62
||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 NAVLELLEDTDVSAIPSEVQGGVWKVFKTPNQDTFLERLNLFLKEGQTVSGMFR 121
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 KVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSRNLFAGEGTVGELS 122
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLL-PMRYRKLRFVSGSAVPAPPEESF 180
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQALQCLKYKLRVFSGRSPSPGEEF 181
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 EVMLEQATEIVKWEW 194
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GRWMFHTTOMIKAW 195
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: September 21, 2004, 13:39:31  
Job time : 46.1859 secs

```

Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQALQCLKYKLRVFSGRSPSPGEEF 181
QY 181 EVMLEQATEIVKWEW 194
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GRWMFHTTOMIKAW 195
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
Q9HAS1 PRELIMINARY; PRT; 351 AA.
AC Q9HAS1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MAP-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=21264738; PubMed=11060313;
RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C.,
RA Yu V.C.;
RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
RT associates with Bax through its Bcl-2 homology domains.";
RL J. Biol. Chem. 276:2802-2807 (2001).
DR EMBL; AF305550; AAG31786.1; -.
SQ SEQUENCE 351 AA; 39521 MW; 8736401AD8A24EAD CRC64;

Query Match 43.8%; Score 436; DB 4; Length 351;
Best Local Similarity 46.9%; Pred. No. 2.9e-31;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

QY 2 LALLEDCRIMSVDEQKSLMTGIPADFEAEIQTQVLQETLKSIGRYLLGKIFRKOENA 61
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LRLLEDWCRCMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGETYLLGRMFRDNR 62
||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 NAVLELLEDTDVSAIPSEVQGGVWKVFKTPNQDTFLERLNLFLKEGQTVSGMFR 121
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 KVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSRNLFAGEGTVGELS 122
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLL-PMRYRKLRFVSGSAVPAPPEESF 180
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQALQCLKYKLRVFSGRSPSPGEEF 181
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 EVMLEQATEIVKWEW 194
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GRWMFHTTOMIKAW 195
:||||| : : : : : : : : : : : : : : : : : : : : : : : : :

```